

REMARKS**Status of the Claims**

Claims 11 and 13-22 are pending in the instant application. Claims 1-10, 12, and 23 have been canceled without prejudice or disclaimer of the subject matter claimed therein. Claims 19-22 have been withdrawn from examination as being drawn to a non-elected invention. Claims 11 and 13-18 are currently under examination.

Applicants thank the Examiner for re-grouping claims 17, 18, and 23 in Group I.

Should the Examiner find claim 17 allowable, we request the opportunity to amend claims 19-22 such that they would be dependent on claim 17 and consequently re-joined as well (see MPEP 821.04).

Amendments to the Claims

Claims 11, 16, and 17 have been amended.

Support for the amendment to claim 11 is found in canceled claim 12.

Support for the amendment to claim 16 is found in Example 1, wherein it shows that the nucleic acid encoding a protein of interest can be integrated into the gene encoding HtrA protease. Thus, the nucleic acid can be integrated, but it need not be integrated into the gene encoding the HtrA protease.

Support for the amendment to claim 17 is found in claim 11 and in claim 17, itself.

These amendments do not introduce prohibited new matter.

Rejection Under 35 U.S.C. § 112, Second Paragraph

Claims 11-18 and 23 have been rejected as being indefinite for failing to particularly point and distinctly claim the subject matter of the invention.

Claims 11, 16, and 23 have been rejected for reciting “HtrA protease” because it is not clear what is meant by the term. As shown on page 1, lines 30 and 31, HtrA protease is a housekeeping serine protease that degrades abnormally or incorrectly folded proteins exported by the bacteria. Pages 1-3 of the specification provide citations that refer to the protease, showing that the protease is well known by its name “HtrA protease.” Additionally, the specification, on page 8, lines 6-11, provides a definition for HtrA protease.

Claims 15 and 18 are rejected for reciting “PrtP protease.” As shown on page 3, lines 34-39, and on page 4, lines 27-35, the term “PrtP protease” refers to a specific protease. Annex 2 and Buist *et al.* indicate that PrtP, a PIII type protease, is structurally and functionally distinct from HtrA protease. Moreover, Applicants have performed a search to show that PrtP and HtrA proteases are distinct proteins. Annex 1 provides the amino acid sequence of HtrA protease of *Lactococcus lactis*. Annex 2 provides the amino acid sequence of PrtP protease of *Lactococcus lactis*. Annex 3 provides a copy of a CDD search performed with the amino acid sequence of HtrA of *Lactococcus lactis* as the query sequence which confirm that PrtP and HtrA proteases are structurally distinct. Accordingly, the PrtP protease in claims 15 and 18 correctly refers to an additional protease.

Claim 13 is rejected for reciting various bacterial strains. Page 7, lines 8-12, and page 9, lines 18-22 provide a list of various Gram positive bacteria that produce the HtrA protease and can be used in the present invention. The preferred embodiment is a *Lactobacillus* strain. However, other Gram positive bacteria express the HtrA protease. The attached result of an internet search (Google) indicate that various bacteria express HtrA protease. Additionally, annexes 4-7, which contain the results of BLAST searches performed against available sequences of several Gram positive bacteria such as *Lactobacilli*, *Lactococci*, and *Streptococci*, confirm that various Gram positive bacteria express the HtrA protease. Thus, the bacterial strains encompassed by claim 13 could be used to practice the method of claim 11.

Accordingly, Applicants respectfully request withdrawal of this rejection.

Rejection Under 35 U.S.C. § 102(b)

A. Claims 23 and 11-15 are rejected under 35 U.S.C. § 102(b) as being anticipated by Bayles *et al.*

Claim 23 has been canceled and claims 11-15 have been amended to recite a step for recovering the protein and that the size of the genome of the bacterial strain is equal to or less than 3.2 Mb.

Bayles *et al.* disclose mutant *Listeria monocytogenes* comprising a *HtrA* gene. However, Bayles *et al.* do not disclose a method of using this mutant bacterium to produce an exported protein. Claims 11-15 are directed to a method of producing a desired protein comprising

culturing a Gram positive bacterial strain that expresses the protein and that has a genome of less than or equal to 3.2 Mb and recovering the protein exported from the bacterial strain.

Accordingly, Bayles *et al.* do not anticipate the claimed invention. Applicants respectfully request withdrawal of this rejection.

B. Claims 23 and 11-15 are rejected under 35 U.S.C. § 102(b) as being anticipated by Buist *et al.*

Claim 23 has been canceled and claims 11-15 have been amended to recite a step for recovering the protein and that the size of the genome of the bacterial strain is equal to or less than 3.2 Mb.

Buist *et al.* disclose a PrtP negative *Lactococcus lactis* strain. However, Buist *et al.* do not disclose a Gram positive bacterial strain that does not express a functional HtrA protease or the use of such a strain to produce a desired protein. As discussed above HtrA protease and PrtP protease are structurally and functionally distinct proteases. Accordingly, Buist *et al.* do not anticipate the claimed invention. Applicants respectfully request withdrawal of this rejection.

The Examiner notes that the prior art teaches multiple protease-deficient strains of *B. subtilis*. Applicants respectfully point out, unlike the bacterial strains used in the claimed invention, *B. subtilis* has a large genome of about 4.2 Mb that encodes several functional HtrA proteases such as YyxA, YkdA, and YvtB/Yirf and numerous other extracellular proteases.

C. Claims 23 and 11-15 are rejected under 35 U.S.C. § 102(b) as being anticipated by Smeds *et al.*

Claim 23 has been canceled and claims 11-15 have been amended to recite a step for recovering the protein and that the size of the genome of the bacterial strain is equal to or less than 3.2 Mb.

Smeds *et al.* disclose a strain of mutant *Lactobacillus helveticus* in which the *gusA* reporter gene was inserted downstream of the *htrA* promoter. The *gusA* reporter gene encodes β-glucuronidase. Although culturing the mutant *Lactobacillus helveticus* induces the *gusA* mRNA, culturing the mutant bacterial strain did not induce the expression of β-glucuronidase, the protein of interest (page 6152 , col.1, second full paragraph). Thus, the cited reference does not and

could not teach recovering the protein of interest. Accordingly, Smeds *et al.* do not anticipate the claimed invention. Applicants respectfully request withdrawal of this rejection.

Rejection Under 35 U.S.C. § 103(a)

Claims 16, 17, and 18 are rejected under 35 U.S.C. § 103(a) as being unpatentable over Bayles *et al.* or Buist *et al.* as applied to claims 23 and 11-15 above, and further in view of any one of Dougan *et al.* or Georgiou *et al.*

Claim 16 is directed to a method of producing a protein of interest using Gram positive bacteria with a genome of less than or equal to 3.2 Mb and that does not express a functional HtrA protease. Claims 17 and 18 are directed to a Gram positive bacterial strain with a genome of less than or equal to 3.2 Mb and that does not express a functional HtrA protease. The small size of the genome leaves no room for the presence of other proteases of the HtrA family or for the presence of other proteases having a similar function. This ensures that no residual proteolytic activity remains after the single HtrA protease has been inactivated.

The deficiencies of Bayles *et al.* and Buist *et al.* are discussed above.

Dougan *et al.* teach Gram-negative bacteria having a mutation in the *degP* gene of the HtrA family for expressing a heterologous antigen. It seems that *degQ* and *degS* genes of the HtrA family are still intact and functional. Thus, the *htrA* protease gene of this bacteria is still functional with respect to the *degQ* and *degS* genes. Accordingly, Dougan *et al.* do not teach the use of Gram positive bacteria that have a genome of less than or equal to 3.2 Mb and that do not express a functional HtrA protease for producing a desired protein.

Similarly, Georgiou *et al.* disclose the use of mutant Gram negative bacteria that are multiply protease deficient for producing proteolytically sensitive polypeptides. Specifically, Georgiou *et al.* teach mutant Gram negative bacteria deficient in DegP, OmpT, and/or Protease III protease. It appears that since the *degQ* and *degS* genes of the HtrA are intact, the *htrA* gene must be functional. Accordingly, Georgiou *et al.* do not teach the use of Gram positive bacteria that have a genome of less than or equal to 3.2 Mb and that do not express a functional HtrA protease, which bacteria is used for producing a desired protein.

Applicants respectfully point out that Georgiou *et al.* stated that inactivation of a single protease is not sufficient to prevent the degradation of exported polypeptides (col. 2, lines 30-

40). Thus, Georgiou *et al.* teach inactivating multiple proteolytic enzymes. However, they point out that there is no assurance that disablement or deletion of any given protease or combination of proteases will result in a viable unchanged host cell or that such manipulation will avoid the precipitation of toxic events within the cell. Accordingly, although it is possible to create mutant organisms having deficiencies in more than three proteases because many Gram negative bacteria express at least seven or eight different proteases that degrade secreted polypeptides (col. 6, 1st paragraph), Georgiou *et al.* stated that deactivating a large number of proteolytic enzymes at some point will compromise the cell's viability (col. 6. 3rd paragraph).

Accordingly, there would not have been any reasonable expectation of success in obtaining the Gram positive bacteria of claims 17 and 18 or the method of using the Gram positive bacteria described in claim 16 by combining the cited references. Furthermore, there would not have been any motivation to combine the cited references since Bayles *et al* and Buist *et al.* teach Gram positive bacteria while Dougan *et al.* and Georgiou *et al.* teach Gram negative bacteria. Gram positive bacteria are different from Gram negative bacteria in many aspects, including production of proteolytic enzymes. Applicants respectfully request withdrawal of the rejection.

Conclusion

The foregoing amendments and remarks are being made to place the application in condition for allowance. Applicants respectfully request entry of the amendments, reconsideration, and the timely allowance of the pending claims. A favorable action is awaited. Should the Examiner find that an interview would be helpful to further prosecution of this application, they are invited to telephone the undersigned at their convenience.

If there are any additional fees due in connection with the filing of this response, please charge the fees to our Deposit Account No. 50-0310. If a fee is required for an extension of time under 37 C.F.R. § 1.136 not accounted for above, such an extension is requested and the fee should also be charged to our Deposit Account.

Respectfully submitted,
Morgan, Lewis & Bockius LLP

Date: November 10, 2003
Morgan, Lewis & Bockius LLP
Customer No. **09629**
1111 Pennsylvania Avenue, N.W.
Washington, D.C. 20004
Tel: 202-739-3000
Fax: 202-739-3001


Sally P. Teng
Registration No. 45,397

NiceProt View of Swiss-Prot: Q9LA06

Entry information

Entry name	HTRA_LACLA
Primary accession number	Q9LA06
Secondary accession numbers	None
Entered in Swiss-Prot in	Release 40, October 2001
Sequence was last modified in	Release 40, October 2001
Annotations were last modified in	Release 41, February 2003

Name and origin of the protein

Protein name	Serine protease do-like htrA	
Synonyms	EC 3.4.21.- HtrAL1	
Gene name	HTRA or LL2136	
From	Lactococcus lactis (subsp. lactis) (Streptococcus lactis)	[TaxID: 1360]
Taxonomy	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.	

References

[1] SEQUENCE FROM NUCLEIC ACID, AND CHARACTERIZATION.
STRAIN=IL1403;
MEDLINE=20177820; PubMed=10712686;
Poquet I., Saint V., Seznec E., Simoes N., Bolotin A., Gruss A.;
"HtrA is the unique surface housekeeping protease in Lactococcus lactis and is required
for natural protein processing.";
Mol. Microbiol. 35:1042-1051(2000).

[2] SEQUENCE FROM NUCLEIC ACID.
STRAIN=IL1403;
MEDLINE=21235186; PubMed=11337471;
Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K., Weissenbach J., Ehrlich S.D.,
Sorokin A.;
"The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssp. lactis
IL1403.";
Genome Res. 11:731-753(2001).

Comments

- **FUNCTION:** DEGRADES ABNORMAL EXPORTED PROTEINS. NEEDED FOR THE PRO-PEPTIDE PROCESSING OF A NATURAL PRO-PROTEIN AND FOR MATURATION OF A NATIVE PROTEIN. RESPONSIBLE FOR THE HOUSEKEEPING OF EXPORTED PROTEINS.
- **SUBCELLULAR LOCATION:** Membrane-bound (*Probable*).
- **SIMILARITY:** Belongs to peptidase family S2C.

ANNEX 1

- **SIMILARITY:** Contains 1 PDZ/DHR domain.

Copyright

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

Cross-references

EMBL	AF155705; AAF61294.1; -. AE006442; AAK06234.1; -.
PIR	H86891; H86891.
MEROPS	S01.273; -.
InterPro	IPR009003; Cys_Ser_trypsin. IPR001478; PDZ. IPR001254; Peptidase_S1. IPR001940; Peptidase_S1C.
Pfam	PF00595; PDZ; 1. PF00089; trypsin; 1.
PRINTS	PR00834; PROTEASES2C.
SMART	SM00228; PDZ; 1.
PROSITE	PS50106; PDZ; 1.
Implicit links to	CMR; ProDom; HOBACGEN; BLOCKS; ProtoNet; ProtoMap; PRESAGE; DIP; ModBase; SWISS-2DPAGE.

Keywords

Hydrolase; Serine protease; Transmembrane; Complete proteome.

Features

Key	From	To	Length	Description
TRANSMEM	6	26	21	POTENTIAL.
DOMAIN	88	284	197	CATALYTIC.
DOMAIN	302	383	82	PDZ.
ACT_SITE	127	127		CHARGE RELAY SYSTEM (POTENTIAL).
ACT_SITE	157	157		CHARGE RELAY SYSTEM (POTENTIAL).
ACT_SITE	239	239		CHARGE RELAY SYSTEM (POTENTIAL).

Sequence information

Length: 408 AA

Molecular weight: 41648 Da

CRC64: 581B90B55A7DF851 [This is a checksum on the sequence]

ANNEX 1

10	20	30	40	50	60
MAKANIGKLL	LTGVVGGAAIA	LGGSAIYQST	TNQSANNRS	NTTSTKVSNV	SVNVNTDVTS
70	80	90	100	110	120
AIKKVSNNSVV	SVMNYQKDNS	QSSDFSSIFG	GNSGSSSTD	GLQLSSEGSG	VIYKKSGGDA
130	140	150	160	170	180
YVVTNYHIVIA	GNSSLVDVLLS	GGQKVKASVV	GYDEYTDLAV	LKISSEHVKD	VATFADSSKL
190	200	210	220	230	240
TIGEPAIAVG	SPLGSQFANT	ATEGILSAT	RQVTLTQENG	QTTNINAIQT	DAAINPGNSG
250	260	270	280	290	300
GALINIEGQV	IGITQSKITT	TEDGSTSVEG	LGFAIPSNDV	VNIINKLEAD	GKISRPALEGI
310	320	330	340	350	360
RMVDLSQLST	NDSSQLKLPS	SVTGGVVVYS	VQSGLPAAASA	GLKAGDVITK	VGDTAVTSST
370	380	390	400		
DLQSALYSHN	INDTVKVTYY	RDGKSNTADV	KLSKSTSDEL	TSSPSSSN	

NiceProt View of Swiss-Prot: P15292

Entry information

Entry name	P3P_LACLC
Primary accession number	P15292
Secondary accession numbers	None
Entered in Swiss-Prot in	Release 14, April 1990
Sequence was last modified in	Release 14, April 1990
Annotations were last modified in	Release 41, February 2003

Name and origin of the protein

Protein name	PIII-type proteinase [Precursor]	
Synonyms	EC 3.4.21.96	
	Lactocepin	
	Cell wall-associated serine proteinase	
Gene name	PRTP	
From	Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris)	[TaxID: 1359]
Encoded on	Plasmid.	
Taxonomy	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.	

References

[1] SEQUENCE FROM NUCLEIC ACID, AND SEQUENCE OF 188-197.

STRAIN=SK11;
 MEDLINE=89340435; PubMed=2760036;
 Vos P., Simons G., Siezen R.J., de Vos W.M.;
 "Primary structure and organization of the gene for a prokaryotic, cell envelope-located
 serine proteinase.";
 J. Biol. Chem. 264:13579-13585(1989).

Comments

- **FUNCTION:** PROTEASE WHICH BREAKS DOWN MILK PROTEINS DURING THE GROWTH OF THE BACTERIA ON MILK.
- **CATALYTIC ACTIVITY:** Endopeptidase activity with very broad specificity, although some subsite preference have been noted, e.g. large hydrophobic residues in the P1 and P4 positions, and Pro in the P2 position. Best known for its action on caseins, although it has been shown to hydrolyze hemoglobin and oxidized insulin B-chain.
- **SUBCELLULAR LOCATION:** Attached to the cell wall peptidoglycan by an amide bond (*Potential*).
- **SIMILARITY:** Belongs to peptidase family S8.

ANNEX 2

Copyright

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch)

Cross-references

EMBL	J04962; AAA03533.1; ALT_SEQ.
HSSP	P00782; 2SBT.
MEROPS	S08.019; -.
InterPro	IPR001899; Gram_pos_anchor. IPR003137; PA. IPR000209; Peptidase_S8.
Pfam	PF00746; Gram_pos_anchor; 1. PF02225; PA; 1. PF00082; Peptidase_S8; 1.
PRINTS	PR00723; SUBTILISIN.
TIGRFAMs	TIGR01167; LPXTG_anchor; 1.
PROSITE	PS50847; GRAM_POS_ANCHORING; 1. PS00136; SUBTILASE ASP; 1. PS00137; SUBTILASE HIS; 1. PS00138; SUBTILASE SER; 1.
Implicit links to	ProDom; HOBACGEN; BLOCKS; ProtoNet; ProtoMap; PRESAGE; DIP; ModBase; SWISS-2DPAGE.

Keywords

Hydrolase; Serine protease; Cell wall; Peptidoglycan-anchor; Zymogen; Signal; Plasmid.

Features

Key	From	To	Length	Description
SIGNAL	1	33	33	
PROPEP	34	187	154	
CHAIN	188	1870	1683	PIII-TYPE PROTEINASE.
PROPEP	1871	1902	32	REMOVED BY SORTASE (POTENTIAL).
ACT_SITE	217	217		CHARGE RELAY SYSTEM (BY SIMILARITY).
ACT_SITE	281	281		CHARGE RELAY SYSTEM (BY SIMILARITY).
ACT_SITE	620	620		CHARGE RELAY SYSTEM (BY SIMILARITY).
SITE	1867	1871	5	LPXTG SORTING SIGNAL (POTENTIAL).
MOD_RES	1870	1870		AMIDE-LINKED TO CELL WALL (POTENTIAL).

ANNEX 2

Sequence information

Length: 1902 AA [This is the length of the unprocessed precursor]

Molecular weight: 200550 Da [This is the MW of the unprocessed precursor]

CRC64: 87CECBAA9345F9D3 [This is a checksum on the sequence]

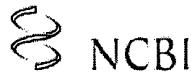
10	20	30	40	50	60
MQRKKKGLSI	LLAGTVALGA	LAVLPVGEIQ	AKAAISQQTK	GSSLANTVTA	ATAKQAATDT
70	80	90	100	110	120
TAATTNQAI	TQLAAKGIDY	NKLNKVQQQD	IYVDVIVQMS	AAPASENGIL	RTDYSSTAEI
130	140	150	160	170	180
QQETNKVIAA	QASVKAABEQ	VTQQTAGESY	GYVVNGFSTK	VRVVVIPKLN	QIAGVKTVTL
190	200	210	220	230	240
AKVYYPTDAK	ANSMANVQAV	WSNYKYKGE	TVVSVIDSGI	DPTHKDMRLS	DDKDVKLTKS
250	260	270	280	290	300
DVEKFTDTV	HGRYFNSKVP	YGFNYADNN	TITDDKVDEQ	HGMHVAGIIG	ANGTGDDPAK
310	320	330	340	350	360
SVVGVAPEAQ	LLAMKVFSNS	DTSAKTGSAT	VVSAIEDSAK	IGADVLNMSL	GSNSGNQTL
370	380	390	400	410	420
DPELAQAVQNA	NESGTAAVIS	AGNSGTSGSA	TEGVNKDYYG	LQDNEMVGSP	GTSRGATTVA
430	440	450	460	470	480
SAENTDVITQ	AVTITDGTGL	QLGPETIQLS	SHDFTGSFDQ	KKFYIVKDAS	GNLSKGALAD
490	500	510	520	530	540
YTADAKGKIA	IVKRGEFSFD	DKQKYAQAAG	AAGLIIVNTD	GTATPMTSIA	LTTTFPTFGL
550	560	570	580	590	600
SSVTGQKLVD	WVTAHPDDSL	GVKITLAMLP	NQKYTEDKMS	DFTSYGPVSN	LSFKPDITAP
610	620	630	640	650	660
GGNIWSTQNN	NGYTNMSGTS	MASPFIAGSQ	ALLKQALNNK	NNPFYAYYKQ	LKGTLTDFL
670	680	690	700	710	720
KTVEMNNTAQP	INDINYNVI	VSPRRQGAGL	VDVKAAIDAL	EKNPSTVVAE	NGYPAVELKD
730	740	750	760	770	780
FTSTDKTFKL	TFTNRRTHEL	TYQMDSNTDT	NAVYTSATDP	NSGVLYDKKI	DGAAIKAGSN
790	800	810	820	830	840

ANNEX 2

ITVPAGKTAQ	IEFTLSLPKS	FDQQQFVEGF	LNFKGSDGSR	LNLPYMGFFG	DWNDGKIVDS
850	860	870	880	890	900
LNGITYSPAG	GNFGTVPLLK	NKNTGTQYYG	GMVTDADGNK	TVDDQAIAFS	SDKNALYNDI
910	920	930	940	950	960
SMKYYLRLNI	SNVQVDILDG	QGNKVTTLSS	STNRKKTYYN	AHSQQYIYYN	APAWDGTYYD
970	980	990	1000	1010	1020
QRDGNIKTAD	DGSYTYRISG	VPEGGDKRQV	FDVPFKLDSK	APTVRHVALS	AKTENGKTQY
1030	1040	1050	1060	1070	1080
YLTAEAKDDL	SGLDATKSVK	TEINEVTNL	ATFTDAGTTA	DGYTKIETPL	SDEQAQALGN
1090	1100	1110	1120	1130	1140
GDNSAELYLT	DNASNATDQD	ASVQKPGSTS	FDLIVNGGGI	PDKISSTTTG	YEANTQGGGT
1150	1160	1170	1180	1190	1200
YTFSGTYPAA	VDGTYTDAQG	KKHDLNTTYD	AATNSFTASM	PVTNADYAAQ	VDLYADKAHT
1210	1220	1230	1240	1250	1260
QLLKHFDTKV	RLMAPFTDLD	KFNNGSDQTS	EATIKVTGTV	SADTKTVNVG	HTVAALDAQH
1270	1280	1290	1300	1310	1320
HFSVDVPVNY	GDNTIKVTAT	DKDGNTTTEQ	KTITSSYDPD	MLKKSVTFDQ	GVKFGTNKFN
1330	1340	1350	1360	1370	1380
ATSAKFYDPK	TGIATITGKV	KHPTTTLQVD	GKQIPIKDDL	TFSFTLDLGT	LGQKPGFVVV
1390	1400	1410	1420	1430	1440
GDTTQNKTTFQ	EALSFILDAV	APTLSLDSST	DAPVYTNDPN	FQITGTATDN	AQYLSLSING
1450	1460	1470	1480	1490	1500
SSVASQYEDI	NINSGKPGHM	AIDQPVKLLE	GKNVLTVAVT	DSEDNTTTKN	ITVYYEPKKT
1510	1520	1530	1540	1550	1560
LAAPTVTPST	TEPAQTVTLT	ANAAATGETV	QYSADGGKTY	QDVPAAGVTI	TANGTFKFKS
1570	1580	1590	1600	1610	1620
TDLYGNESPA	VDYVVVTNIKA	DDPAQLQAAK	QELTNLIASA	KTLSASGKYD	DATTTALAAA
1630	1640	1650	1660	1670	1680
TQKAQTALDQ	TNASVDSLTD	ANRDLQTAIN	QLAAKLPADK	KTSLLNQLQS	VKDALGTDLG
1690	1700	1710	1720	1730	1740

ANNEX 2

NQTDPSTGKT FTAALDDLVA QAQAGTQTDD QLQATLAKIL DEVLAKLAEG IKAATPAEVG
1750 1760 1770 1780 1790 1800
NAKDAATGKT WYADIADTLT SGQASADASD KLAHLQALQS LTKVAAAVE AAKTVGKGDG
1810 1820 1830 1840 1850 1860
TTGTSDKGGS QGTPAPAPGD TGKDKGDEGS QPSSGGNIPT KPATTTSTTT DDTTDRNGQL
1870 1880 1890 1900
TSGKGALPKT GETTERPAFG FLGVIVVSLM GVLGLKRKQR EE



Conserved Domain Databases

		PubMed	Nucleotide	Protein	Structure	CDD	Taxonomy	Help?					
CD: COG0265.1, DegQ, Query added				PSSM-Id: 10140		Source: Cog							
Description: Trypsin-like serine proteases, typically periplasmic, contain C-terminal PDZ domain [Posttranslational modification, protein turnover, chaperones]													
Taxa: cellular organisms			Related: may span multiple domains										
Status: Alignment from source			Created: 7-Oct-2002										
Aligned: 135 rows			PSSM: 347 columns			Representative: Consensus							
Proteins: [Click here for CDART summary of Proteins containing COG0265]													
						10 20 30 40 50 60							
					*....*....*....*....*....*....							
consensus	1	LLVLAGLDLAVG-----		LLLIAIAIGG-----		RALTSAA-- 29							
query	1			makanigkllltgvvggaiyalggssaiyq		s 29							
gi 15616584	22	IGISAFIGAILGal-----	lvLFSPALSGLgwlp	e-----	dsgapteETGQLTEap	70							
gi 22001651	47	WFRPLLGVIGGsla-----	lgiyTFTPLGNHDsqdtakq-----	sssqqqtQSVTATSts	97								
gi 1731364	23	YFLSSLIGVIVGav-----	lmaFIMPYLsNEgl	dtg-----	aldqQ--QNNNgr	64							
gi 14194653	6	igkLLLTVGGGaia-----	lggsAIYQSTTNQs-----		aNNSRNSNTs 44								
gi 15840667	141	AAAALGTPALAApaphgalagsgk	gLGVRDVLFGKvsylal	gilvaialviGGIGGVigr	200								
gi 15902042	14	LLVVIVISFFSGal-----	gsFSITQLTQKSS-----		vNNNSNNNSti 50								
gi 15675945	10	sLSILLIGFLGGli-----	aiITFNNLYPHSp-----		SKINSGKAtt 46								
gi 16799397	122	YFLTALIGVIIGgl-----	iIFFVAWDNGDnad-----		ttS--NSNNkp 158								
			70 80 90 100 110 120		*....*....*....*....*....							
consensus	30	-----GQR-----	LSFATAVEKVAPAVVSIATGLTAKL-----		R 58								
query	30	ttnqsannrsnttstkvsnvsvnvntdvt	tsaiKKvsnsvvsvmnyqkdnsqssdf	ssfif	89								
gi 15616584	71	ndietvnYAVn-----	SDVSQAVEKVSDAVV-GIVSMTNGs-----		105								
gi 22001651	98	seskkssSSSsafk-----	sedssSKIISDMVEDLSPAIV-GITNLQAQSns-----		141								
gi 1731364	65	esirtvnvsvn-----	NAVTKIVSNMSPAVV-GVVMNIQKsd-----		99								
gi 14194653	45	tkvsnvsVNVn-----	TDVTSIAKKVSNNSV-SVMNYQKDNsqssd-fssifG	90									
gi 15840667	201	ktaevvdAFTtskv1sttgnaQEPA	GRTKVAAGAVADSVVTIESV-----		246								
gi 15902042	51	t---qtaYKNe-----	NSTTQAVNKVKDAVV-SVITYSANRqns-----		vfg 88								
gi 15675945	47	s---nmvFNNT-----	TNTTKAVKAVQNAV-SVINYQDNPssslsnpytklfG	91									
gi 16799397	159	tkvekvsvNTT-----	SDVTKAVIDKVQDAVV-SVLYNQSSSs 1-----		195								
			130 140 150 160 170 180		*....*....*....*....*....							
consensus	59	SFF-PSDPP--LRSAEGLGSGFIIS-----	SDG-----YIVTNNHVIAAGAEITVTL-----		102								
query	90	ggnsgsssstdgqlqlssEGSGVIYK-----	KSGgdaYVVTNHYVIAGNSSLDVLL-----		139								
gi 15616584	106	-mFsSSEE-----	EGTGSQVIYKkegDRA-----FIVTNEHVISGANQVe-----		146								
gi 22001651	142	s1FgSSSSDssEDTESGSGSVIFKkenGKA-----	YIITNNHVVEGASSLkv-----		189								
gi 1731364	100	-iWgESGE-----	AGSGSGSVIYKkndHSA-----YVVTNHHVIEGASQIei-----		139								
gi 14194653	91	GNSgSSSStdgLQ-LSSEGSQVIYKksgGDA-----	YVVTNHYVIAGNSSLdv-----		137								
gi 15840667	247	-----SDQE-----	GMQGSGSVIVD-GRG-----YIVTNNHVISEAANN-----		286								
gi 15902042	89	Nd--DDTDDsQ-----	ISSEGSGSVIYKkndKEA-----YIVTNNHVINGASKVdi-----		132								
gi 15675945	92	EGRsKENKDaeLS-IFSEGSGSVIYRkdgNSA-----	YVVTNHHVIDGAKRIei-----		138								
gi 16799397	196	dgTtSSEKE-----	ASSGSGSVIYKkanGKA-----YIVTNNHVADANKLev-----		237								
			190 200 210 220 230 240		*....*....*....*....*....							
consensus	103	--ADGREVPALKVGKDPISDLAVLKIDGAGGLPVIALGDSDKLVRGVVVVAIGNPFG-----	L 158										
query	140	--SGGQKVKA SVVGYDEYT-----	DLAVLKISSEHVKDVATFADSSKLTI	GEPAIAVGSP	L GsqF 197								
gi 15616584	147	v1TDGSRLPAEVLGSDVFT-----	DLAVLEIDGSDVETVAEFGNSDLLSPGEPAIAIGNPLGk	lF 206									
gi 22001651	190	s1YDGTEVTAKLGVGSDSLTD-----	LAVLQISDDHVTKVANFGDSSLRGTGETVIAIGDPLGk	lF 249									
gi 1731364	140	s1KDGSRSVADLVGSDQLMD-----	LA LRVKSDKIKAVADFGNSDKVKS	GEPAIAIGNPLGleF 199									
gi 14194653	138	11SGGQKVKA SVVGYDEYT-----	DLAVLKISSEHVKDVATFADSSKLTI	GEPAIAVGSP	L GsqF 197								
gi 15840667	287	vfNDGKEVPANLVGRDPK-----	DLAVLKVDNVNDNLTVARLGDSKVRGDEVLAVGAPLG-----	L 344									
gi 15902042	133	r1SDGTVKPGGEIVGADTFSDIA-----	VKISSEKVTTAEFGDSSKLTVGETAIAIGSPLGseY	192									
gi 15675945	139	1mADGSKVVGELVGADTYSL-----	DAVVKISSDKIKTVAEFADSTKLN	VGEVIAIGSPLGtqY 198									
gi 16799397	238	tftNGKSEAKLLG-----	DLAVLEIDDKNVSTVATFGDSDLKLGEPAIAIGSPLGteF	297									

ANNEX 3

		250	260	270	280	290	300	
	*.....*.....*.....*.....*.....*.....*.....						
consensus	159	GQTVTSGIVSALGR---T-GVGSAGG---YVNFIQTDAAINPGNSGGPLVNIDGEVV						208
query	198	ANTATEGILSATSRqvlt1T-QENGQTT---NINAIQTDAAINPGNSGGALINIEGQVI						251
<u>gi_15616584</u>	207	<u>SSSVTGLIISATER---SiPIDLTGNqidwQAEVLQTDAAINPGNSGGALVNQGQVI</u>						262
<u>gi_22001651</u>	250	<u>SRTVTQGIVSGVDR---TvSMSTSAGE---tSINVQTDAAINPGNSGGPLNTDGKIV</u>						302
<u>gi_1731364</u>	200	<u>AGSVTQGVISGTER---AiPVDSNGDgqpdwNAEVLQTDAAINPGNSGGALLNMDGKVI</u>						255
<u>gi_14194653</u>	198	<u>ANTATEGILSATSR---QvTLTQENGqt---tNINAIQTDAAINPGNSGGALINIEGQVI</u>						251
<u>gi_15840667</u>	345	<u>RSTVTQGIVSALHR---PvPLSGEGSdt-dtVIDAIQTDASINHGNSGGPLIDMDAQVI</u>						399
<u>gi_15902042</u>	193	<u>ANTVTQGIVSSLNR---NvSLKSEDGqa---iSTKAIQTDAAINPGNSGGPLINIQGQVI</u>						246
<u>gi_15675945</u>	199	<u>ANSVTQGIVSSLR---TvTLKNENGet--vSTNAIQTDAAINPGNSGGPLINIEGQVI</u>						252
<u>gi_16799397</u>	298	<u>SGSVTQGIIISGLNR---AvPVDTNGDgtedwEADVIQTDAAINPGNSGGALINIEGQVI</u>						353
		310	320	330	340	350	360	
	*.....*.....*.....*.....*.....*.....						
consensus	209	<u>GIN---TAIAPSGG-----SSGIGFAIPVNLVAPVLDLISKGVVRGYLGVIGE</u>						256
query	252	<u>GITqskittTEDGSTS-----VEGLGFAIPSNDVVNIINKLEADGKISRPAVGIRMV</u>						303
<u>gi_15616584</u>	263	<u>GIN---SMKIAQSt-----VEGIGFAIPSNLAIPIVIEDLEFYGDVQRPQMVGAFR</u>						309
<u>gi_22001651</u>	303	<u>GIN---SMKISEDD-----VEGIGFAIPSNDVKPIAEELLSKGQIERPYIGVSM</u>						349
<u>gi_1731364</u>	256	<u>GIN---SMKIAESA-----VEGIGLIPSPLKLVIPVIEDLERYGKVKRPFGLIEMK</u>						302
<u>gi_14194653</u>	252	<u>GIT---QSKITTTEdgs---tsVEGLGFAIPSNDVVNIINKLEADGKISRPAVGIRMV</u>						303
<u>gi_15840667</u>	400	<u>GIN---TAGKSLsd---sasGLGFAIPVNEMKLVANSLIKDGKIVHPTLGISTR</u>						447
<u>gi_15902042</u>	247	<u>GIT---SSKIATnng---tsVEGLGFAIPANDAINIIEQLEKNGKTRPALGIQMV</u>						296
<u>gi_15675945</u>	253	<u>GIN---SSKISSTPTgsngnsgaVEGIGFAIPSTDVIKIIQKLETNGEVIRPALGISMV</u>						308
<u>gi_16799397</u>	354	<u>GIN---SMKISMEN-----VEGISFAIPSNTVEPIIEQLETKGEVERPSLGVSLR</u>						400
		370	380	390	400	410	420	
	*.....*.....*.....*.....*.....						
consensus	257	<u>PLTADIA----LGLP---VAAGAVVLGVLPGS-----PAAKAGIKAGDIITAVNGK</u>						300
query	304	<u>DLSQLST----NDSS---QLKLPSVTGGVVVYsvqsglPAASAGLKAGDVITKVGD</u>						354
<u>gi_15616584</u>	310	<u>SLSEIPsfhweetLKLPe-dvKGGVSVTIDIVPM斯-----PAETAGLRQYDVIVELNGE</u>						361
<u>gi_22001651</u>	350	<u>DLEQVPQnyqegtLGLFgsqLNKGVYIREVASGS-----PAEKAGLKAEDIIGLKGK</u>						402
<u>gi_1731364</u>	303	<u>SLSDIASyhwdetLKLPk-nVTNGAVVMGVDafs-----PAGKAGLKELDVITEFDGY</u>						354
<u>gi_14194653</u>	304	<u>DLSQLSTnds-sqLKLPs-sVTGGVVVYSVQSGL-----PAASAGLKAGDVITKVGD</u>						354
<u>gi_15840667</u>	448	<u>--S---VSn-----aIASGAQVANVKAGS-----PAQKGGILENDVIVKVGNR</u>						485
<u>gi_15902042</u>	297	<u>NLSNVSTsdi-rrLNIPs-nVTSGIVRSVQSNN-----PAN-GHLEYKDVTKVDDK</u>						346
<u>gi_15675945</u>	309	<u>NLNDLStnal-sqINIPT-sVTGGIVVAEVKEGM-----PAS-GKLAQYDVITEIDGK</u>						358
<u>gi_16799397</u>	401	<u>DVDTIPetqqknLKLp-sVDYGMVQQVVSGS-----AADKAGLKQYDVIVELNGE</u>						452
		430	440	450	460			
	*.....*.....*.....*.....*.....						
consensus	301	<u>PVASLSDLVAAVASNR--PGDEVALKLLRGGKERELAVTLGdrSPLSAS</u>						347
query	355	<u>AVTSSTDQLQSALYSHN--INDTVKVITYYRDGKSNTADVklksstsdslet</u>						401
<u>gi_15616584</u>	362	<u>DINDGHELRKFLYTELn-IGDEVEVTYYREGKKETTTLTL---VEQQSS</u>						406
<u>gi_22001651</u>	403	<u>EIDTGSELRNILYKDAK-IGDTVEVKILRNGKEMTKKIKLDQ-KEEKTS</u>						449
<u>gi_1731364</u>	355	<u>KVNDIVDLRKRLYQKK--VGDRVVKFYRGGKEKSVDIKLSS-ADQLGS</u>						400
<u>gi_14194653</u>	355	<u>AVTSSTDQLQSALYSHN--INDTVKVITYYRDGKSNTADVklksstsdslet</u>						400
<u>gi_15840667</u>	486	<u>AVADSDEFVVAVRQLA---IGQDAPIEVVREGRHVTLTVKPD-PST---</u>						528
<u>gi_15902042</u>	347	<u>EIASSTDQLQSALYNHS---IGDTIKITYRNGKEETTSIKLNK-SSGDLE</u>						392
<u>gi_15675945</u>	359	<u>TVNSISDLQSSLYGHD--INDTIKVTYRGTTKKKADIKLTK-TTQDLT</u>						404
<u>gi_16799397</u>	453	<u>KVTNSMTLRKILYGNdvkIGDKVKYRDGKEQSTDIKLEA-AKTTT-</u>						499

ANNEX 4

BLASTP 2.2.6 [Apr-09-2003]

RID: 1067437801-292-1030821.BLASTQ3

Query= gi|15674118|ref|NP_268293.1| exported serine protease

[Lactococcus lactis subsp. lactis]

(408 letters)

Database: Unfinished *Lactobacillus gasseri*; Completed *Lactobacillus plantarum* WCFS1;

Completed *Lactococcus lactis* subsp. *lactis*;

Unfinished *Leuconostoc mesenteroides* subsp. *mesenteroides* ATCC 8293;

Unfinished *Oenococcus oeni* MCW;

Completed *Streptococcus agalactiae* 2603V/R;

Completed *Streptococcus agalactiae* NEM316;

15,229 sequences; 4,501,851 total letters

Taxonomy reports

Score (bits)	E Value
Sequences producing significant alignments:	
ref NP_268293.1 exported serine protease [Lactococcus lact...	576 e-165
ref NP_689159.1 serine protease [Streptococcus agalactiae ...	303 5e-83
ref NP_783901.1 serine protease HtrA [Lactobacillus planta...	283 5e-77
ref ZP_00063134.1 COG0265: Trypsin-like serine proteases, ...	281 2e-76
ref ZP_00069121.1 COG0265: Trypsin-like serine proteases, ...	272 7e-74
ref ZP_00046803.1 COG0265: Trypsin-like serine proteases, ...	230 3e-61
ref ZP_00070364.1 COG0265: Trypsin-like serine proteases, ...	181 2e-46
ref ZP_00064063.1 COG0265: Trypsin-like serine proteases, ...	180 4e-46
ref ZP_00070156.1 COG0750: Predicted membrane-associated Z...	45 2e-05
ref NP_266705.1 UDP-N-acetylglucosamine 1-carboxyvinyltran...	35 0.034
ref ZP_00046513.1 COG2996: Uncharacterized protein conserv...	33 0.097
ref ZP_00063264.1 COG0750: Predicted membrane-associated Z...	33 0.13
ref NP_785411.1 carboxy-terminal processing proteinase [La...	33 0.13
ref NP_786668.1 extracellular protein [Lactobacillus plant...	32 0.28
ref NP_268285.1 hypothetical protein [Lactococcus lactis s...	31 0.48
ref NP_267651.1 sugar ABC transporter substrate binding pr...	30 1.1
ref NP_687067.1 peptidase, M23/M37 family [Streptococcus a...	30 1.1

ANNEX 4

ref NP_784951.1	cell surface SD repeat protein precursor [...]	29	1.4
ref NP_786644.1	extracellular protein, gamma-D-glutamate-m... methyltransferase [putative] [Lactobacillus plantarum]	29	1.4
ref NP_687090.1	alcohol dehydrogenase, propanol-preferring... [putative] [Lactococcus lactis ssp. lactis]	29	1.8
ref NP_267008.1	hypothetical protein [Lactococcus lactis s...] [putative] [Lactococcus lactis ssp. lactis]	29	1.8
ref NP_734524.1	Unknown [Streptococcus agalactiae NEM316]	29	1.8
ref ZP_00064050.1	COG1364: N-acetylglutamate synthase (N-a...) [putative] [Lactococcus lactis ssp. lactis]	28	2.4
ref ZP_00063238.1	COG1674: DNA segregation ATPase FtsK/Spo... [putative] [Lactococcus lactis ssp. lactis]	28	3.1
ref NP_735868.1	Unknown [Streptococcus agalactiae NEM316]	28	3.1
ref NP_784552.1	acetyltransferase (putative) [Lactobacillu...] [putative] [Lactococcus lactis ssp. lactis]	28	3.1
ref ZP_00070200.1	COG1477: Membrane-associated lipoprotein... [putative] [Lactococcus lactis ssp. lactis]	28	3.1
ref ZP_00063415.1	COG1668: ABC-type Na ⁺ efflux pump, perme... [putative] [Lactobacillus plantarum]	28	4.1
ref NP_785643.1	endopeptidase La (putative) [Lactobacillus...] [putative] [Lactococcus lactis ssp. lactis]	28	4.1
ref ZP_00069981.1	COG3051: Citrate lyase, alpha subunit [O...] [putative] [Lactococcus lactis ssp. lactis]	28	4.1
ref ZP_00064376.1	COG1364: N-acetylglutamate synthase (N-a...) [putative] [Lactococcus lactis ssp. lactis]	28	4.1
ref NP_688903.1	membrane-associated zinc metalloprotease, ... [putative] [Lactococcus lactis ssp. lactis]	28	4.1
ref NP_268318.1	hypothetical protein [Lactococcus lactis s...] [putative] [Lactococcus lactis ssp. lactis]	27	5.3
ref ZP_00046283.1	COG0507: ATP-dependent exoDNase (exonucl...) [putative] [Lactococcus lactis ssp. lactis]	27	5.3
ref ZP_00069420.1	COG3480: Predicted secreted protein cont... [putative] [Lactococcus lactis ssp. lactis]	27	5.3
ref ZP_00063200.1	COG0827: Adenine-specific DNA methylase ... [putative] [Lactococcus lactis ssp. lactis]	27	7.0
ref ZP_00062802.1	hypothetical protein [Leuconostoc mesent...] [putative] [Leuconostoc mesenteroides ssp. me...]	27	7.0
ref NP_687818.1	major facilitator family protein [Streptoc...] [putative] [Streptococcus agalactiae NEM316]	27	7.0
ref NP_688359.1	conserved hypothetical protein [Streptococ...] [putative] [Streptococcus agalactiae NEM316]	27	7.0
ref NP_688028.1	sensor histidine kinase, putative [Strepto...] [putative] [Streptococcus agalactiae NEM316]	27	7.0
ref NP_735272.1	Unknown [Streptococcus agalactiae NEM316]	27	7.0
ref NP_786635.1	extracellular protein [Lactobacillus plant...] [putative] [Lactococcus lactis ssp. lactis]	27	7.0
ref ZP_00046678.1	COG4653: Predicted phage phi-C31 gp36 ma... [putative] [Lactococcus lactis ssp. lactis]	27	7.0
ref ZP_00046947.1	COG2931: RTX toxins and related Ca ²⁺ -bin... [putative] [Lactococcus lactis ssp. lactis]	27	7.0
ref ZP_00046780.1	COG3210: Large exoproteins involved in h... [putative] [Lactococcus lactis ssp. lactis]	27	7.0
ref ZP_00062638.1	COG4932: Predicted outer membrane protei... [putative] [Lactococcus lactis ssp. lactis]	27	9.1
ref NP_687888.1	exonuclease RexA [Streptococcus agalactiae...] [putative] [Streptococcus agalactiae NEM316]	27	9.1
ref NP_687383.1	3-oxoacyl-(acyl-carrier-protein) synthase ... [putative] [Streptococcus agalactiae NEM316]	27	9.1

Alignments

>ref|NP_268293.1| exported serine protease [Lactococcus lactis subsp. lactis]
Length = 408

Score = 576 bits (1484), Expect = e-165
Identities = 310/390 (79%), Positives = 310/390 (79%)

```

Query: 1  MAKANIGKLLLTGVVGGAIALGGSAYQXXXXXXXXXXXXXXXXXXXXXX 60
        MAKANIGKLLLTGVVGGAIALGGSAYQ
Sbjct: 1  MAKANIGKLLLTGVVGGAIALGGSAYQSTTNQSANNRSNTTSTKVSNVNVNTDVTS 60

Query: 61  AIKKXXXXXXXXXNMYQKDNSQXXXXXXXXXXXXXXXXXXXXEGSGVYKKSGGDA 120
          AIKK          MNYQKDNSQ          EGSGVYKKSGGDA
Sbjct: 61  AIKKVSNSVSVMVNYQKDNSQSSDFSSIFGGNSGSSSTDGLQLSEGSGVYKKSGGDA 120

Query: 121  YVVTNYHVIAGNSSLDVLLSGGQKVAKASVVGYDEYTDLAVLKISSEHVKDVATFADSSKL 180
          YVVTNYHVIAGNSSLDVLLSGGQKVAKASVVGYDEYTDLAVLKISSEHVKDVATFADSSKL
Sbjct: 121  YVVTNYHVIAGNSSLDVLLSGGQKVAKASVVGYDEYTDLAVLKISSEHVKDVATFADSSKL 180

Query: 181  TIGEPAIAVGSPPLGSQFANTATEGILSATSRQVTLTQENGQTTNNAIQTDAAINPGNSG 240
          TIGEPAIAVGSPPLGSQFANTATEGILSATSRQVTLTQENGQTTNNAIQTDAAINPGNSG
Sbjct: 181  TIGEPAIAVGSPPLGSQFANTATEGILSATSRQVTLTQENGQTTNNAIQTDAAINPGNSG 240

Query: 241  GALINIEGQVIGITQSKITTTEGSTSVEGLGFAIPSNDVNNIINKLEADGKISRPA 300
          GALINIEGQVIGITQSKITTTEGSTSVEGLGFAIPSNDVNNIINKLEADGKISRPA
Sbjct: 241  GALINIEGQVIGITQSKITTTEGSTSVEGLGFAIPSNDVNNIINKLEADGKISRPA 300

Query: 301  RMVDLSQLSTNDSQLKLPXXXXXXXXXXXXXXLPAASAGLKAGDVITKGDTAVTSST 360

```

ANNEX 4

RMVDLSQLSTNDSSQLKLP LPAASAGLKAGDVITKVGDTAVTSST
Sbjct: 301 RMVDLSQLSTNDSSQLKLPSSVTGGVVVSVQSGLPAASAGLKAGDVITKVGDTAVTSST 360

Query: 361 DLQSALYSHNINDTVKVTVYYRDGKSNTADV 390
DLQSALYSHNINDTVKVTVYYRDGKSNTADV
Sbjct: 361 DLQSALYSHNINDTVKVTVYYRDGKSNTADV 390

>ref|NP_689159.1| serine protease [Streptococcus agalactiae 2603V/R]
ref|NP_736563.1| Unknown [Streptococcus agalactiae NEM316]
Length = 409

Score = 303 bits (775), Expect = 5e-83
Identities = 160/290 (55%), Positives = 201/290 (69%), Gaps = 7/290 (2%)

Query: 107 EGSGVIYKKSGGDAYVVTNYHVIAGNSSLVLLSGGQKVKA
Sbjct: 105 EGSGVIYKKDGKNA
Length = 166

EGSGVIYKK G +AYVVTN HV G + + L+ G K +VG D Y+DLAV+KI S+
Sbjct: 105 EGSGVIYKKDGKNA
Length = 164

Query: 167 HVKDVATFADSSKLTIGEPAIAVGSP
Sbjct: 165 KVSNIAEFADSSKLNIGET
Length = 226

V ++A FADSSKL IGE AIA+GSPLG++AN+ T+GI+S+ R VT+T E GQT + N
Sbjct: 165 KVSNIAEFADSSKLNIGET
Length = 224

Query: 227 AIQTDAAINPGNSGGALINIEGOVIGITQSKITT
Sbjct: 225 AIQTDAAINPGNSGGALINIEGOVIGINSS
Length = 280

AIQTDAAINPGNSGGALINIEGOVIGI SKI++T + SVEG+GFAIPSNDV
Sbjct: 225 AIQTDAAINPGNSGGALINIEGOVIGINSS
Length = 284

Query: 281 VNIINKLEADGKISRPALGIRMV
Sbjct: 285 VKIINQLESNGQVERPALGISMAGLSNLPSDV
Length = 340

V IIN+LE++G++ RPALGI M LS L ++ S+LK+P +P A
Sbjct: 285 VKIINQLESNGQVERPALGISMAGLSNLPSDV
Length = 343

Query: 341 GLKAGDVITKVGDTAVTSSTD
Sbjct: 344 KLKKYDVITKVDDKEVVSPSDLQ
Length = 390

LK DVITKV D V S +DLQS LY H + D++ VT+YR T +
Sbjct: 344 KLKKYDVITKVDDKEVVSPSDLQ
Length = 393

>ref|NP_783901.1| serine protease HtrA [Lactobacillus plantarum WCFS1]
Length = 420

Score = 283 bits (723), Expect = 5e-77
Identities = 162/394 (41%), Positives = 214/394 (54%), Gaps = 14/394 (3%)

Query: 10 LLTGVVGGAI
Sbjct: 14 LVAGLIGGGVAYGGINYFQNNNIAT
Length = 65

ALGGSAIYQXXXXXXXXXXXXXXXXXXXXXX---XXXXXXAIKKX
L+ G++GG +A GG +Q +
Sbjct: 14 LVAGLIGGGVAYGGINYFQNNNIAT
Length = 73

Query: 66 XXXXXXXMNYQKDNSQ
Sbjct: 74 KAAVVSVINLQK
Length = 118

KDNXXXXXXX---XXXXXXXXXXXXXXXXXXXXXXEGSGV
+N QK +S EGSG+IYKKSG
Sbjct: 74 KAAVVSVINLQK
Length = 133

Query: 119 DAYVVTNYHVIAGNSSLVLLSGGQKVKA
Sbjct: 134 AAYIVVNNHVGSSAIRVIMSDGT
Length = 178

QVIGITQSKITT
AY+VTN HV++G+S++ V++S G K+ A +VG D TDEAVLKI+S V A+F +S
Sbjct: 134 AAYIVVNNHVGSSAIRVIMSDGT
Length = 193

Query: 179 KLTIGEPAIAVGSP
Sbjct: 194 NIKVGETAL
Length = 237

LSQFANTATEGILSATSRQV
+ +GE A+A+GSP+GS +A T T+GI+SA R V T +GQTT IQTD AIN G
Sbjct: 194 NIKVGETAL
Length = 253

Query: 238 NSGGALINIEGOVIGITQSKITT
Sbjct: 254 NSGGPLFNIAGOVIGINS
Length = 297

EDGSTSVEGLGFAIPSNDVV
NSGG ALNI GOVIGI K+ + G TSVEG+GFAIPS
Sbjct: 254 NSGGPLFNIAGOVIGINS
Length = 312

ANNEX 4

Query: 298 LGIRMVDLSQLSTND-SSQLKLPXXXXXXXXXXXXXXLPAASAGLKAGDVITKVGDTAV 356
LG+ DLS +S++D S LKLP PA +AGL DVIT++G V
Sbjct: 313 LGVATYDLSNISSSDQKSVLKLPTSVTKGVVIMKTYSGSPAAGLTKYDVITELGKKV 372

Query: 357 TSSTDLQSALYSHNINDTVKVTVYYRDGKSNTADV 390
TS L+SALY+H++NDTV V YY +GK TA++
Sbjct: 373 TSLATLRSALYAHSVNDTWTVKYHNGKLKTANM 406

>ref|ZP_00063134.1| COG0265: Trypsin-like serine proteases, typically periplasmic,
contain C-terminal PDZ domain [Leuconostoc mesenteroides
subsp. mesenteroides ATCC 8293]
Length = 379

Score = 281 bits (718), Expect = 2e-76
Identities = 160/392 (40%), Positives = 214/392 (54%), Gaps = 23/392 (5%)

Query: 1 MAKANIGKLLLTVGGAIALGGSAYQXXXXXXXXXXXXXXXXXXXXXX 60
M + + K LLTGV+ G + GG+ +Y
Sbjct: 1 MVQPALTKEEGLGGV-GGAILYQQGVQLQNQKVSTTATSTKTIAKNATATS 59

Query: 61 AIKKXXXXXXXXXMYQKDNSQXXXXXXXXXXXXXXXXXXXXEGSGVIVKKSCGDA 120
A K +N+ K + EGSGVIVKK+ G A
Sbjct: 60 AYNKVSADAVSVLNFTKSSQ-----GSYQESSEGSVIVKKTDGSA 100

Query: 121 YVVVTNYHVIAGNSSLDVLLSGGQKVKA SVVGDEYTDIAV LKISSEHVKVATFADSSKL 180
++VTVHVI G + + V+L G+KV A++VG D TDI AVLKI V A F DSSK+
Sbjct: 101 EIVVTVNHVITGAAKIQVMLHSGKKVTATLVGKDAMTDLAVL KIDGTDVTTAQFGDSSKI 160

Query: 181 TIGEPAIAVGSPPLGSQFANTATEGILSATSRQVTLTQENGQT-TNINAIQTDAAINPGNS 239
T+GE +A+GSPLGS++A++ T+GI+SA R V T ENQ IQTDAAINPGNS
Sbjct: 161 TVGENVLAIGSPLGSEYASSVTQGIISAKKRLVEATSENGQNYGGSTVIQTDAAINPGNS 220

Query: 240 GGALLINIEGQVIGITQSKitTTEDGTSVEGLGFAIPSNDVVIINKLEADGKISRPAIG 299
GG-LTNGQVIGI K++T+ G TSVEG+GFAIPS+ VV+I+NKL DGK++RPA+G
Sbjct: 221 GGPLLINFAGQVIGINSMKLSTSSSG-TSVEGMGFAIPSDQVVDIVNKLVKGKVTRPAIG 279

Query: 300 IRMVSQLSTND-SSQLKLPXXXXXXXXXXXXXXLPAASAGLKAGDVITKVGDTAVTS 358
I +++LS+++ ++ S LK+P PA AGLK DVI + V+S
Sbjct: 280 ISLINLSEVTASEQKSTLKPDSVTGGVVMSLTNNNGPADKAGLKKYDVIVGINGKKVSS 339

Query: 359 STDLQSALYSHNINDTVKVTVYYRDGKSNTADV 390
DL+ LY H++ DT+ +TYY T V
Sbjct: 340 QADLREELYKHSLGDTITLTYYHQDTKQTVKV 371

>ref|ZP_00069121.1| COG0265: Trypsin-like serine proteases, typically periplasmic,
contain C-terminal PDZ domain [Oenococcus oeni MCW]
Length = 425

Score = 272 bits (696), Expect = 7e-74
Identities = 160/390 (41%), Positives = 210/390 (53%), Gaps = 6/390 (1%)

Query: 6 IGKLLLTVGGAIALGGSAY-QXXXXXXXXXXXXXXXXXXXXAIKK 64
I LL G++GG +A+G IY Q
Sbjct: 29 IATALLAGLLGGGVAVGAGYIYTQTTDFIGKSTGALSDGKTTIKAPTISGKSATKVYNN 88

Query: 65 XXXXXXXXMYQKDNSQXXXXXXXXXXXXXXXXXXXXEGSGVIVKKSCGDA YVVT 124

ANNEX 4

+N Q +S EGSGVIYK + G AYVVT
Sbjct: 89 LKGAVVSVINQQATSSSTIYGDSKKSSSTSSFSTLQTASEGSGVIYKDADGYAYVVT 148

Query: 125 NYHVIAGNSSLDVLLSGGQKVKA SVVG YDEYTDLA VLKISSEHVKD VATFADSSKL TIGE 184
NYHVI+G + V+L GG KV A VG D TDLA VL+IS VK VA F +S+++ G+
Sbjct: 149 NYHVISGAKRIQVVLYGGTKVVA KVGSDAMTDLA VL RISGSDVKTVAQFGNSNQIKTGQ 208

Query: 185 PAIAVGSPPLGSQFANTATEGILSATSRQVTLTQENGQTT--NINA QTDAAINPGNSGG 242
+A+GSPLG+ +A++ TEGI+SA+ R V+ T E+G+T + AI QTDAAINPGNSGG
Sbjct: 209 TVLAIGSPPLGTDYASSVTEGIISASKRLVSNTSES GKTNYGDSIAI QTDAAINPGNSGG 268

Query: 243 LINIEGOVIGITQS KITTTEDGSTSVEGLGFAIPSNDVVNIINKLEADGKISRPA LGIRM 302
L+N+ GOVIGI K+T T++G SVEG+GFAIPSN VV+IINKL GK+ RPALG+ +
Sbjct: 269 LVNTSGOVIGINSQKL TETDEGE-SVEGMGFAIPSNTVVSIIINKLIK YGKVV RPALGVEV 327

Query: 303 VDLSQLSTN-DSSQLKL PX-XXXXXXXXXXXXXXLPAASAGL KAGDVITKVGDTAVTSST 360
VDLS++S++ LKLP PA AG+K DVI V V++
Sbjct: 328 VDLSEVSSDVVKKTLKLPSKVKTGIVIAGFSSDKSPAKKAGIKKYDVIVAVNGEKVSNLA 387

Query: 361 DLQSALYSHNINDTVKV TYYRDGKSNTADV 390
D++ +Y + DTVK+TYYR T V
Sbjct: 388 DMRDIYKLKVGD TVKITYYRASTEKTVKV 417

>ref|ZP_00046803.1| COG0265: Trypsin-like serine proteases, typically periplasmic,
contain C-terminal PDZ domain [Lactobacillus gasseri]
Length = 666

Score = 230 bits (587), Expect = 3e-61
Identities = 123/243 (50%), Positives = 162/243 (66%), Gaps = 2/243 (0%)

Query: 107 EGSGVIYKKSGGDAYVVTNYHVIAGNSSLDVLLSGGQKVKA SVVG YDEYTDLA VLKISSE 166
EGSGVIY KS G V+VTN HV++G+ + V+LS G+KV A VG D TDLA VL I +
Sbjct: 133 EGSGVIYMKNSNGKGVIVTNNHVVSGSDEIQVILSNGKKVTA KVGTDSETDLA VL TIDGK 192

Query: 167 HVKDVA TFADSSKLTIGEPAIAVGSPPLGSQFANTATEGILSATSRQVTLTQENGQTTN-I 225
+V A F S L G+ IAVGSPLG S++A + T+GI+SA +R V +T GQ TN
Sbjct: 193 YVTQTAQFGSSKNLEPGQQVIAVGSPLGSEYATSVTQGIISAKNRTVDVTNSAGQVTNQA 252

Query: 226 NAI QTDAAINPGNSGGALINIEGOVIGITQS KITTTEDGSTSVEGLGFAIPSNDVVNIIN 285
IQTDAAINPGNSGG L+N+ GOVIGI K+ +DG T+VEG+GFAIP S++VV+IIN
Sbjct: 253 TVI QTDAAINPGNSGGPLVNMSGOVIGINSMKLSSSDG-TAVEGMGFAIPSDEVVSIIN 311

Query: 286 KLEADGKISRPA LGIRMVDLSQLSTNDSQLKL PXXXXXXXXXXXXXLPAA SAGL KAG 345
+L +GK I+RP LG+R+V + +L +L LP A AG+K+
Sbjct: 312 QLVKNGKIRPKLGV RVVSVDEL TEYGRKKLGPDSVKSGVYVASVTKNGSADKAGIKSH 371

Query: 346 DVI 348
DVI
Sbjct: 372 DVI 374

>ref|ZP_00070364.1| COG0265: Trypsin-like serine proteases, typically periplasmic,
contain C-terminal PDZ domain [Oenococcus oeni MCW]
Length = 301

Score = 181 bits (460), Expect = 2e-46
Identities = 100/285 (35%), Positives = 148/285 (51%), Gaps = 12/285 (4%)

ANNEX 4

Query: 10 LLTGVVGGAIALGGSAYQXXXXXXXXXXXXXXXXXXXX--XXXXXXXXXXAIKKXXX 67
LL+ ++G + LG ++ A K
Sbjct: 15 LLSAIIAGATVVLGCFYLFYLAPAQNKAAKSSSIAAGMTKVVNLGTSSSQATKAYNKVKN 74

Query: 68 XXXXXMNYQKDNSQXXXX----XXXXXXXXXXXXXXXXXXXXEGSGVIVKKSGGDAYV 122
NYQK +++ EG+G+IY+ G +Y+
Sbjct: 75 AVVTVENYQKPSTEASDYFFEWFGSQSGGSSSSSTEDQLAAEGTGLIYETDGNYSYI 134

Query: 123 VTNHVIAGNSSLDVLLSGGQKVAKASVVGYDEYTDLAVLKISSEHVKDVFADSSKLT 182
VTN HVI G + +++++ G KVKA ++G + D+AVL+ISS V TF +SSK+
Sbjct: 135 VTNNHVIKGANEIEIIMANGTKVKAKLIGKNATKDIAVRLRISSASVTTGTFVNSSKVQA 194

Query: 183 GEPAIAVGSPPLGSQFANTATEGILSATSRQVTLTQENGQTTNINAIQTDAAINPGNSGGA 242
G+ +A+GSPLGS +A++ T GI+SAT+RQ+ + +AIQTD A+NPGNSGG
Sbjct: 195 GQQVLAIGSPLGSDYASSLTSGIVSATNRQI----DDSPIKLSAIQTDVALNPGNSGGP 249

Query: 243 LINIEGGVIGITQSKITTTEDGSTSVEGLGFAIPSNDVVNIINKL 287
LIN+ G+VIGI KI++TEDGS VEG+ F+IPSN VV I +
Sbjct: 250 LINMAGEVIGINSMKISSTEDGSEDVEGMSFSIPSNTVVATIKSI 294

>ref|ZP_00064063.1| COG0265: Trypsin-like serine proteases, typically periplasmic,
contain C-terminal PDZ domain [Leuconostoc mesenteroides
subsp. mesenteroides ATCC 8293]
Length = 253

Score = 180 bits (457), Expect = 4e-46
Identities = 93/182 (51%), Positives = 129/182 (70%), Gaps = 3/182 (1%)

Query: 107 EGSGVYIKKSGGDAYVVTNHYVIAGNSSLDVLLSGGQKVAKASVVGYDEYTDLAVLKISSE 166
EGSGV+YK SGG AYV TN+IV+A + L ++ + G+K++A++VG D DLATLK +
Sbjct: 74 EGSGVVYKISGGYAYVLTNNHVADSDELQLITASGKKIQATIVGTDSSKDLALLKAKTT 133

Query: 167 HVKDVFATFADSSKLTIGEPAIAVGSPPLGSQFANTATEGILSATSRQVTLTQENGQTTNIN 226
+K A+F ++ KL G+ +A+GSPLGS +A + T GI+SA R TL+ E ++
Sbjct: 134 DIKTSASFNAKKLQSGQQVLAIGSPLGSDYATSLTGIVSAPRR--TLSAETGSSATT 191

Query: 227 AIQTDAAINPGNSGGALINIEGGVIGITQSKITTTEDGSTSVEGLGFAIPSNDVVNIINK 286
AIQTDAAINPGNSGG LIN+GQV+GI SKI ++ DG TSVEG+GFAIP++ V I
Sbjct: 192 AIQTDAAINPGNSGGPLINLKQGVVGINSSKIASSTDG-TSVEGMGFAIPADIVQTFIKN 250

Query: 287 LE 288
E
Sbjct: 251 TE 252

>ref|ZP_00070156.1| COG0750: Predicted membrane-associated Zn-dependent proteases 1
[Oenococcus oeni MCW]
Length = 421

Score = 45.1 bits (105), Expect = 2e-05
Identities = 26/56 (46%), Positives = 35/56 (62%), Gaps = 3/56 (5%)

Query: 336 PAASAGLKAGDVITKVGDITAVTSSTDLQSALYSHNIND-TVKVTVYYRDGKSNTADV 390
PA GLK GDVITKV + +++ T L +A+ N+ D T+KV+Y R KS T V
Sbjct: 218 PAMKQGLKKGDVITKVDSLKISNWTQLTTAI--ENVGDKTMKVSYRRGNKSRTVTV 271

>ref|NP_266705.1| UDP-N-acetylglucosamine 1-carboxyvinyltransferase [Lactococcus

ANNEX 4

lactis subsp. lactis]
Length = 427

Score = 34.7 bits (78), Expect = 0.034
Identities = 40/153 (26%), Positives = 71/153 (46%), Gaps = 14/153 (9%)

Query: 148 SVVGYDEYTDALVKISSEHVKDVTADSSKLTIGEPAIAVGSPLGSQF--ANTATEGI 205
+ + +D+ + K +SE +K A + SK+ +I V P+ ++ A + G
Sbjct: 65 TAISFDQEAKKIIAKSNSE-IKTTAPYEYVSKM--RASIVVMGPILARNGQARVSMPGG 120

Query: 206 LSATSRQVTLT---QENGQTTNINAIQTDAAINPGNSGGALINIEGQVIGITQSKI--T 259
S SR + L ++ G T NA +A + GA I ++ +G TQ+ I
Sbjct: 121 CSIGSRPIDLHLRGFEQMGATITQNAGYIEAKAD--KLKGAIHYLDFPSVGATQNLILAA 178

Query: 260 TTEDGSTSVEGLGFAIPSNDVVNIINKLEADGK 292
T DG+T++E D+ N++NK+ A+ K
Sbjct: 179 TLADGTTLENAAREPEIVDLANLLNKGANVK 211

>ref|ZP_00046513.1| COG2996: Uncharacterized protein conserved in bacteria
[Lactobacillus gasseri]
Length = 297

Score = 33.1 bits (74), Expect = 0.097
Identities = 22/68 (32%), Positives = 37/68 (54%), Gaps = 4/68 (5%)

Query: 109 SGVIYKSGGDAYVVTNYHvia--GNSSLDVDLLSGGQKVKA SVVGYDEY--TDLAVLKIS 164
SG +Y+ ++V+T+ + +A S + L GQK+KA V+G +Y +L+VL
Sbjct: 157 SGTVYRNVEVGSFVITDQYYLAFVHKSEMFRPLRLGQKIKARVIGVSQYGRNLNSVLPNG 216

Query: 165 SEHVKDVA 172
E + D A
Sbjct: 217 FEEIDDDA 224

>ref|ZP_00063264.1| COG0750: Predicted membrane-associated Zn-dependent
proteases 1
[Leuconostoc mesenteroides subsp. mesenteroides ATCC
8293]
Length = 417

Score = 32.7 bits (73), Expect = 0.13
Identities = 18/56 (32%), Positives = 30/56 (53%), Gaps = 1/56 (1%)

Query: 335 LPAASAGLKAGDVITKVGDTAVTSSTDLQSALYSHNINDTVKVTVYYRDGKSNTADV 390
+PA AGLKAGD IT++ D T++ D + ++ + +T R+G +V
Sbjct: 212 MPADQAGLKAGDEITQI-DRVKTWTWDQVANAIGNSKESQLNITVLRNGHKQVEV 266

>ref|NP_785411.1| carboxy-terminal processing proteinase [Lactobacillus
plantarum
WCFS1]
Length = 492

Score = 32.7 bits (73), Expect = 0.13
Identities = 20/56 (35%), Positives = 29/56 (51%), Gaps = 1/56 (1%)

Query: 336 PAASAGLKAGDVITKVGDTAVTSSTDLQS-ALYSHNINDTVKVTVYYRDGKSNTADV 390
PA AGLK D+I V +V T Q+ ++ I TVK+T R G++ T +
Sbjct: 147 PAKKAGLKPKDIIKAVNGKSVAGKLTQAVSMMRGKIGTTVKLTIERSGQTFTVSL 202

ANNEX 4

>ref|NP_786668.1| extracellular protein [Lactobacillus plantarum WCFS1]
Length = 190

Score = 31.6 bits (70), Expect = 0.28
Identities = 32/141 (22%), Positives = 55/141 (39%), Gaps = 32/141 (22%)

Query: 187 IAVGSPLGSQFANTAT-----EGILSATSRQVTLTQENGQTTNINAIQTDA----- 232
+ G PL Q A+T T E I T++ +TL Q G + I D+
Sbjct: 14 LMAGLPLVGQAADTETTKAEEVELIQQDDTNKDITLDQAPGVSGTEKITNDSTYDAKNV 73

Query: 233 -----AINPGNSGGALINIEGQVI-----GITQSKITTEDGSTSV-EGLGFA 274
NPGN+ G L+ ++G +T +++ T D + ++ + +
Sbjct: 74 TGDLKVTPGNTDGWLVQVKGSKFMNADDTRELRAALTFAQVNATADDANNISKAKAYK 133

Query: 275 IPSNDVVNIINKLEADGKISR 295
+ D II EA+ I +
Sbjct: 134 VDITDQNQIIMDAEANESEGIGK 154

>ref|NP_268285.1| hypothetical protein [Lactococcus lactis subsp. lactis]
Length = 428

Score = 30.8 bits (68), Expect = 0.48
Identities = 20/55 (36%), Positives = 28/55 (50%), Gaps = 1/55 (1%)

Query: 336 PAASAGLKAGDVITKVGDTAVTSSTDQLQSALYSHNINDTVKVTVYYRDGKSNTADV 390
PA +AGLKAGD I V T ++ + S + +K+ R GKS T V
Sbjct: 225 PAYNAGLKAGDKIEAVNGTKTADWNNVVTEI-SGSKGKELKLEVSRSGKSETLSV 278

>ref|NP_267651.1| sugar ABC transporter substrate binding protein [Lactococcus lactis subsp. lactis]
Length = 483

Score = 29.6 bits (65), Expect = 1.1
Identities = 20/77 (25%), Positives = 35/77 (45%), Gaps = 1/77 (1%)

Query: 125 NYHVIAGNSSLVDVLLSGGQKVKASVVGYDEYTDLAVLKISSEHVKDVTAFADSSKLIGE 184
NY + N++ + G K+ S +G+ +Y + +SS D+A FA +
Sbjct: 49 NYKELMANANKILEKKAGVKLDISYIGWGDYAQKMNIVSSGEAYDIA-FAQDYATNAAK 107

Query: 185 PAIAVGSPPLGSQFANTA 201
A A + L ++A TA
Sbjct: 108 GAFADLTDLAPKYAKTA 124

>ref|NP_687067.1| peptidase, M23/M37 family [Streptococcus agalactiae 2603V/R]
>ref|NP_734500.1| Unknown [Streptococcus agalactiae NEM316]
Length = 299

Score = 29.6 bits (65), Expect = 1.1
Identities = 27/134 (20%), Positives = 50/134 (37%), Gaps = 7/134 (5%)

Query: 239 SGGALINIEGQVIGITQSKITTEDGSTSVEGLGFAIPSNDVVNIINKLEADGKISRPA 298
S G+ + + V I +ITT +G G+ +A+P+ ++ + ADG +
Sbjct: 20 SAGSRVLADTYVRPIDNRTTGNGYPGHCGVDYAVPTGTIIRAV---ADGTVKFAGA 75

Query: 299 GIRMVDLSQLSTNDSQLKLPXXXXXXXXXXXXXXLPAASAGLKAGDVITKVGDTAVTS 358
G ++ L+ N + + + +K GD+I VG T + +

ANNEX 4

Sbjct: 76 GANFSWMTDLAGN---CVMIQHADGMHSGYAHMSRVVARTGEVKQGDIIGYVGATGMAT 132

Query: 359 STDLQSALYSHNIN 372

L N N

Sbjct: 133 GPHLHFELPANPN 146

>ref|NP_784951.1| cell surface SD repeat protein precursor [Lactobacillus plantarum WCFS1]

Length = 3360

Score = 29.3 bits (64), Expect = 1.4

Identities = 42/145 (28%), Positives = 57/145 (39%), Gaps = 17/145 (11%)

Query: 164 SSEHVKDVTAFADSSKLTIGEPAIAVGSPPLGS-----QFANTATEGILSATSRQVTLT 216

S + V + SS LT+ A GS L AN T ++ +V +

Sbjct: 1167 SYDAVDSSAGLLSTSSLTVTIKAGYTGSLLFQAVQGFSWDLANWFTVYTFASNLAEVDVY 1226

Query: 217 QENGQTTNINAIQTDAAINPGN-SGGALINIEGQVIGITQ----SKIT-TTEDGS-TSV 268

N TNI+ D INP N S G+ + Q T KIT TT D S ++

Sbjct: 1227 SSNIPIATNISIAGDDYVINPTNSSGSNDKVTSQFTTTNPENATGKITWTTSDDSSIATI 1286

Query: 269 EGLG-FAIPSNDVVNIINKL-EADG 291

+ G + SN V I + ADG

Sbjct: 1287 DDSGLLTUVSNGTVTITATITNAADG 1311

>ref|NP_786644.1| extracellular protein, gamma-D-glutamate-meso-diaminopimelate muropeptidase (putative) [Lactobacillus plantarum WCFS1]

Length = 370

Score = 29.3 bits (64), Expect = 1.4

Identities = 29/111 (26%), Positives = 42/111 (37%), Gaps = 3/111 (2%)

Query: 132 NSSLDVLLSGGQKVAKASVVGDEYTDLAVLKISSEHVKDVTAFADSSKLTIGEPAIAVGS 191

+SS+ S AS V T + SS V AT S+ + A +

Sbjct: 144 SSSVAAQSSSTSTASASSVTSSASTSSVASQASSAVTSSATSQSSASQSSASQASQSST 203

Query: 192 PLGSQFANTATEGILSATSRQVTLTQENGQTTNINAIQTDAAINPGNSGGA 242

P+ S + TAT +ATS T +Q + +N + T A S A

Sbjct: 204 PVASSTTTATSTQSAATS---TSSQASSTASNTSSTTATATAYSASA 251

>ref|NP_687090.1| alcohol dehydrogenase, propanol-preferring [Streptococcus agalactiae 2603V/R]

Length = 338

Score = 28.9 bits (63), Expect = 1.8

Identities = 19/69 (27%), Positives = 36/69 (52%), Gaps = 3/69 (4%)

Query: 112 IYKKSGG-DAYVVTNYHVIAGNSSLVLLSGGQKVAKASVVGDEYTDLAVLKISSEHVKD 170

I +K+GG VVT +A N ++D + +GG V + EY +L+++K + ++

Sbjct: 224 IQEKTGGCHGVVTAWSKVAFNQAIDSVRAGGTVVAVGLP--SEYMELSIVKTVLDGIRV 281

Query: 171 VATFADSSK 179

V + + K

Sbjct: 282 VGSLVGTRK 290

>ref|NP_267008.1| hypothetical protein [Lactococcus lactis subsp. lactis]

ANNEX 4

Length = 1063

Score = 28.9 bits (63), Expect = 1.8
Identities = 53/272 (19%), Positives = 99/272 (36%), Gaps = 30/272 (11%)

Query: 120 AYVVNTNYHVIAGNSSLVLLSGGQKVKASVVGYDEYTDALVKISSEHVKDVFADSSK 179
A ++ + + N D + Q++ G E T + L S+ DVA A ++
Sbjct: 34 AIIIVSGTITDQNVKADTAIDSSQGIS---GITEVTSYSALASSTN--SDVA--ASQNQ 85

Query: 180 LTIGEPAIAVGSPPLGSQFANTATEGILSATSRQVLTQENGQTTNINAIQTDAAINPGNS 239
+ + + + + T TEGI S S E+ TT+ IQT P N+
Sbjct: 86 VAYEQASDQSSNKSLANTVETDTEGITSNVSDSSNSINESQNTTSTVVIQT----PTNN 140

Query: 240 GGALINIEGQVIGITQSKitTEDGSTSVEGLGFAIPSND-VVNIINKLEADGKISRPA 298
++ + S ++ GS S+ A S D V + + G + S +
Sbjct: 141 -----IVSLADSS-SSNDNGSNSILSSSNAADSVDSA VGSQSSTSSGVLESS- 188

Query: 299 GIRMVDSLSQLSTNDSSQLKLPXXXXXXXXXXXXXXLPAASAGLKAGDVITKVGDTAVTS 358
+D S + SS++ L + +T+ A +
Sbjct: 189 ---AIDSGIASVSQSSEMNLVGNSSAASSAAVASFTAILATNPSMVPMLTQALAAAAPA 245

Query: 359 STDLQSLALYSHNINDTVKVTVYYRDGKSNTADV 390
+T SA+ + + D V G S A++
Sbjct: 246 TTS-GSAILNTTLGDLVNQAISTVGISGLANI 276

>ref|NP_734524.1| Unknown [Streptococcus agalactiae NEM316]
Length = 338

Score = 28.9 bits (63), Expect = 1.8
Identities = 19/69 (27%), Positives = 36/69 (52%), Gaps = 3/69 (4%)

Query: 112 IYKKSGG-DAYVVTNYHVIAGNSSLVLLSGGQKVKASVVGYDEYTDALVKISSEHVKD 170
I +K+GG VVT +A N +D + +GG V + EY +L+++K + ++
Sbjct: 224 IQEKTCGGCHGVVVTAVSKVAFNQAI DSVRAGGTVVAVGLP--SEYMELSIVKTVDGIRV 281

Query: 171 VATFADSSK 179
V + + K
Sbjct: 282 VGSLVGRTRK 290

>ref|ZP_00064050.1| COG1364: N-acetylglutamate synthase (N-acetylmornithine aminotransferase) [Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293]
Length = 344

Score = 28.5 bits (62), Expect = 2.4
Identities = 38/195 (19%), Positives = 77/195 (39%), Gaps = 22/195 (11%)

Query: 204 GILSATSQRVLTQENGQTTNINAIQTDAAINPGNSSGALINI-----EGQVIGITQSK 257
G+ + Q Q + +T +Q + N GN+ +I Q Q
Sbjct: 51 GVFTTNLVQAAPVQLDKKTIRNGQLQA-II VNSGNANAVTGSIGVSHAESMQEFTAQQLN 109

Query: 258 ITTTEDGSTSVEGLGFAIPSNDVVIINKLEADGKISRPA LGIRMVDSLSQLSTNDSSQLK 317
I T+ G S +G +P + ++N I +L+ DG + A I D + S S ++
Sbjct: 110 IDTSLVGVASTGIIGKVLPIDKIINGIKQLKIDGDTNGFAHAIMTTDTKEKSITIQSTIQ 169

Query: 318 LPXXXXXXXXXXXXXXLPAASAGLKAGDVITKVGDTAVTSSTDLSALYSHNINDTVKV 377
A +G+ ++ T +G +T+ ++ + L +++ V+
Sbjct: 170 GKIVTMSGV-----AKGSGMLHPNMA TMLG--FITT DINIDAKLLQQALSEDVET 217

ANNEX 4

Query: 378 TYYR---DGKSNTAD 389
++ + DG ++T D
Sbjct: 218 SFNQITIDGDTSTND 232

>ref|ZP_00063238.1| COG1674: DNA segregation ATPase FtsK/SpoIIIE and related proteins
[Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293]
Length = 368

Score = 28.1 bits (61), Expect = 3.1
Identities = 29/108 (26%), Positives = 42/108 (38%), Gaps = 7/108 (6%)

Query: 194 GSQFANTATEGILSATSRQVTLTQENGQTTNINAIQTDAAINPGNSGGALINIEGQVIGI 253
G+ +NT I+ Q + + T TD I NS A N + ++
Sbjct: 253 GAFISNTDVTNIVEFVKSQQEYQSDAMTV-----TDEEIAQDNSENADGNSDELFOE 306

Query: 254 TQSKITTEDGSTSVEGLGFAIPSNDVVNIINKLEADGKISRPAIGR 301
+ + STS+ F I N +I+ LEA G I PA G R
Sbjct: 307 ALQFVIEQQKASTSLLQRRFRIGYNRAARLIDDLEAGGYIG-PADGSR 353

>ref|NP_735868.1| Unknown [Streptococcus agalactiae NEM316]
Length = 414

Score = 28.1 bits (61), Expect = 3.1
Identities = 35/165 (21%), Positives = 60/165 (36%), Gaps = 19/165 (11%)

Query: 137 VLLSGGQKVAKAS--VVGYDEYTDLAVLKISSEHVKDVATFADSSKLTIPEAIAVGSPLG 194
V + G K+ A +V YD T A ++ + VA ++ K T PA+ +
Sbjct: 86 VTVKVGDKITAGQQQLVQYDTTTAQAAYDTANRQLNKVARQINNLKTTGSLPAMELSDQSS 145

Query: 195 SQFANTATEGILSATSRQVLTQE-----NGQTTNINAIQTDAAINPGNSGGALINIEG 248
S T+ AT+R Q N Q ++N DA + AL
Sbjct: 146 SSSQQGQGTQSTSGATNRLLQQNYQSQANASYNQQLQDLNDAYADAQAEVNKAQKAL---- 200

Query: 249 QVIGITQSKitTEDGSTSVEGLGFAIPSNDVVNIINKLEADGKI 293
+ T++ T VE P++ ++ + +GK+
Sbjct: 201 -----NDTVITSVDVSGTVVEVNSIDPASKTSQVLVHVATEGKL 239

>ref|NP_784552.1| acetyltransferase (putative) [Lactobacillus plantarum WCFS1]
Length = 171

Score = 28.1 bits (61), Expect = 3.1
Identities = 28/72 (38%), Positives = 40/72 (55%), Gaps = 13/72 (18%)

Query: 165 SEHVKDVA-TFADSSKLTIPEAIAVGSPLSQFANTAT--EGI--LSATSRSQVLTQE 218
+E V DV A AD+++L +A+ + LG + +NT T EGI LS T Q + +
Sbjct: 2 AEEVVDVRPAEVADAQL-----LALLAQLGRE-SNTFTVDEGIEDLSETDEQAQIERI 54

Query: 219 NGQTTNINAIQT 230
NG TTNI + T
Sbjct: 55 NGTTTNIIFVAT 66

>ref|ZP_00070200.1| COG1477: Membrane-associated lipoprotein involved in thiamine biosynthesis [Oenococcus oeni MCW]
Length = 358

ANNEX 4

Score = 28.1 bits (61), Expect = 3.1
Identities = 12/45 (26%), Positives = 24/45 (53%)

Query: 117 GGDAYVVTNYHVIAGNSSLVDVLLSGGQKVKASVVGYDEYTDLAVL 161
GG+ YV+ H +G +V + + + S VGY +D+++
Sbjct: 210 GGNIYVIGKSHPTSGTRDWNVGIQNPQNQSRGSSVGVRESDMSIV 254

>ref|ZP_00063415.1| COG1668: ABC-type Na⁺ efflux pump, permease component
[Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293]
Length = 438

Score = 27.7 bits (60), Expect = 4.1
Identities = 27/111 (24%), Positives = 47/111 (42%), Gaps = 8/111 (7%)

Query: 131 GNSSLVLLSGGQKVKASVVGYDEYTDLAVLKISSEHVKDVFAD--SSKLTIGE-PAI 187
GN++ ++ + G Q+V++ +V ++ D+ V I++E + A + LT+ + A
Sbjct: 80 GNTTPNIAVGNQEVRSLVQSEKELDIHSNITNEKKANTALQNEKLDGVLTVNKNEAT 139

Query: 188 AVGSPLGSQLFANTATEGILSATSRQVTLTQENGQTTNINAIQTDAAINPGN 238
P Q IL SR TQ + A QT + P N
Sbjct: 140 ITTQPKEKITAIGNLRSQKATQ----YGLTAEQTADLVQPYN 185

>ref|NP_785643.1| endopeptidase La (putative) [Lactobacillus plantarum WCFS1]
Length = 348

Score = 27.7 bits (60), Expect = 4.1
Identities = 14/42 (33%), Positives = 20/42 (47%)

Query: 342 LKAGDVITKVGDTAVTSSTDLQSLAYSHNINDTVKVTVYYRDG 383
LK GD ITKV +++ Q + + V +TY R G
Sbjct: 149 LKVGDITKVGHHFNTASAYQHYIGKQGVGHRVTITYRRKG 190

>ref|ZP_00069981.1| COG3051: Citrate lyase, alpha subunit [Oenococcus oeni MCW]
Length = 449

Score = 27.7 bits (60), Expect = 4.1
Identities = 24/112 (21%), Positives = 45/112 (40%), Gaps = 4/112 (3%)

Query: 193 LGSQFANTATEGILSATSRQVTLTQENGQTTNINAIQTDAAINPGNSGGALINIEGQVIG 252
LG + A + + + V ++G TNI + + + S G L N V
Sbjct: 26 LGIKDLTLAPSSLTNVMNDMVVKAIKSGTITNITSSGMRGSLGDAVSHGLKN--PVVFR 83

Query: 253 ITQSKitTEDGSTSVEGLGFAIPSNDDVVNIINKLEADGKISRPALEGIMVD 304
++ E+G ++ +P++D V N+E D +LG + +D
Sbjct: 84 SHGNRARAIEEGKIKIDVAFLGPNSDEVGNANGMEGDAAFG--SLGYALMD 133

>ref|ZP_00064376.1| COG1364: N-acetylglutamate synthase (N-acetylmethionine aminotransferase) [Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293]
Length = 346

Score = 27.7 bits (60), Expect = 4.1
Identities = 28/138 (20%), Positives = 59/138 (42%), Gaps = 15/138 (10%)

Query: 255 QSKITTEDGSTSVEGLGFAIPSNDDVVNIINKLEADGKISRPALEGIMVDLSQLSTNDSS 314

ANNEX 4

Q I T+ G S +G +P + ++N I +L+ DG + A I D + S S
Sbjct: 49 QLNIDTSLVGVASTGIIGKVLPIDKIINGIKQLKIDGDTNGFAHAIMTTDTKEKSITIQS 108

Query: 315 QLKLPXXXXXXXXXXXXXXPAASAGLKAGDVITKVGDATAVTSSTDLQSALYSHNINDT 374
++ A +G+ ++ T+G +T+ ++ + L +++
Sbjct: 109 TIQGKIVTMSGV-----AKGSGMLHPNMATMLG--FITTDINIDAKLLQQALSED 156

Query: 375 VKVTYYR---DGKSNTAD 389
V+ ++ + DG ++T D
Sbjct: 157 VETSFNQITIDGDTSTND 174

>ref|NP_688903.1| membrane-associated zinc metalloprotease, putative
[Streptococcus
agalactiae 2603V/R]
ref|NP_736335.1| Unknown [Streptococcus agalactiae NEM316]
Length = 419

Score = 27.7 bits (60), Expect = 4.1
Identities = 14/31 (45%), Positives = 20/31 (64%)

Query: 336 PAASAGLKAGDVITKVGDATAVTSSTDLQSAL 366
PAASAGLK D I ++G V++ L +A+
Sbjct: 212 PAASAGLKNNDRILQIGSHKVSNWEQLTAAC 242

>ref|NP_268318.1| hypothetical protein [Lactococcus lactis subsp. lactis]
Length = 342

Score = 27.3 bits (59), Expect = 5.3
Identities = 13/40 (32%), Positives = 20/40 (50%)

Query: 342 LKAGDVITKVGDATAVTSSTDLQSALYSHNINDTVKVVTYYR 381
L+ D IT V TSS D+ + + + D+V + Y R
Sbjct: 151 LELADTITAVNGQQFTSSADMIAYVSKQKVGSVTIEYTR 190

>ref|ZP_00046283.1| COG0507: ATP-dependent exoDNase (exonuclease V), alpha
subunit -
helicase superfamily I member [Lactobacillus gasseri]
Length = 792

Score = 27.3 bits (59), Expect = 5.3
Identities = 24/96 (25%), Positives = 43/96 (44%), Gaps = 11/96 (11%)

Query: 233 AINPGNSGGALIN---IEGQVIGITQSKITT-----EDGSTSVEGLGFAIPSNDVV 281
A N G + G +N + G ++ I QS ++T +D T L A +D+
Sbjct: 201 ADNIGQALGIELNDPKRVRGAILSILQSALSTLGDTYVALDDLLTQAYDLVQASSYDDLA 260

Query: 282 NIINKLEADGKISRPALGIRMVSQLSTNDSSQLK 317
N +N+L+ GK+ + + Q + S++LK
Sbjct: 261 NSVNELQRQGKVVVSGDKAALQGIFQTELDISNELK 296

>ref|ZP_00069420.1| COG3480: Predicted secreted protein containing a PDZ domain
[Oenococcus oeni MCW]
Length = 364

Score = 27.3 bits (59), Expect = 5.3
Identities = 14/44 (31%), Positives = 22/44 (50%)

ANNEX 4

Query: 342 LKAGDVITKVGDATAVTSSTDLQSALYSHNINDTVKVTYYRDGKS 385
+K GD ITKV +S Q L + + + V +T R+ K+
Sbjct: 153 IKVGDTITKVDGKHFNNSAGYQKYLAAMPVGEKVTLTVRNNKT 196

>ref|ZP_00063200.1| COG0827: Adenine-specific DNA methylase [Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293]
Length = 329

Score = 26.9 bits (58), Expect = 7.0
Identities = 20/66 (30%), Positives = 30/66 (45%), Gaps = 1/66 (1%)

Query: 251 IGITQSKitTEDGSTSVEGLFAIPSNDVVNII-NKLEADGKISRPAvgIRMVdLSQLS 309
I + I ED ++ F PSNDVV II + ++ D + PA + + L+ L
Sbjct: 26 ISYIDALIEILEDINSQTvhREFDKPSNDVVQIIQSTIDMDWSLLSPAERKRALQALVLK 85

Query: 310 TNDSSQ 315
N Q
Sbjct: 86 ANREDQ 91

>ref|ZP_00062802.1| hypothetical protein [Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293]
Length = 179

Score = 26.9 bits (58), Expect = 7.0
Identities = 28/135 (20%), Positives = 52/135 (38%), Gaps = 5/135 (3%)

Query: 128 VIAGNSSLDVDLLSGGQKVKA SVVGYDEYTDLAVLKISSEHVKDVATFADSSKL TIGEPAI 187
+ L L S GQ + A+ + + T + V + + + K + + K G +
Sbjct: 7 LFGSEKKLSQLKSTGQ-INATRLARNNDTPVLVAPVTGDLQKITDSRDEPFKTKNGVMLV 65

Query: 188 AVGSPLGSQFANTATEGILSATSRQVTLTQENGQTTNINAIQTDAAINPGNSGGALINIE 247
L + + TE +T+ +TLT + QT + + T + G
Sbjct: 66 PHSGNLMAPVSGIVTE---STNDYLTLDISEQTVTVVGTNSVVRLAQYGVGQQQLHA 121

Query: 248 GQVIGITQSKITTTE 262
G VIG T K+ + +
Sbjct: 122 GDVIGTTNQKVLSAD 136

>ref|NP_687818.1| major facilitator family protein [Streptococcus agalactiae 2603V/R]
Length = 383

Score = 26.9 bits (58), Expect = 7.0
Identities = 13/26 (50%), Positives = 17/26 (65%)

Query: 2 AKANIGKLLLGVVGGAI ALGGSAIY 27
A NIGK L T +VG +A+G + IY
Sbjct: 141 ASLNIGKALTFIVGLVLAIGVNYIY 166

>ref|NP_688359.1| conserved hypothetical protein [Streptococcus agalactiae 2603V/R]
Length = 414

Score = 26.9 bits (58), Expect = 7.0
Identities = 35/165 (21%), Positives = 59/165 (35%), Gaps = 19/165 (11%)

ANNEX 4

Query: 137 VLLSGGQKVAKAS--VVGYDEYTDLAVLKISSEHVKDVTAFADSSKLTIGEPAIAVGSPLG 194
V + G K+ A +V YD T A ++ + VA ++ K T PA+
Sbjct: 86 VTVKVGDKITAGQQLVQYDTTTAQAAAYDTANRQLNKVARQINNLKTTGSLPAMESSDQSS 145

Query: 195 SQFANTATEGILSATSRQVTLTQE-----NGQTTNINAIQTDAAINPGNSGGALINIEG 248
S T+ AT+R Q N Q ++N DA + AL
Sbjct: 146 SSSQGQGTQSTSGATNRLLQQNYQSQANASYNQQQLQDLNDAYADAQAEVNKAQKAL---- 200

Query: 249 QVIGITQSKITTTEGSTSVEGLGFAIPSNDVNNIINKLEADGKI 293
+ T++ T VE P++ ++ + +GK+
Sbjct: 201 -----NDTVITSDVSGTVVEVNSDIDPASKTSQVLVHVATEGKL 239

>ref|NP_688028.1| sensor histidine kinase, putative [Streptococcus agalactiae 2603V/R]
ref|NP_735501.1| Unknown [Streptococcus agalactiae NEM316]
Length = 579

Score = 26.9 bits (58), Expect = 7.0
Identities = 12/47 (25%), Positives = 28/47 (59%), Gaps = 1/47 (2%)

Query: 147 ASVVGYDEYTDLAVLKISSEHVKDVTAFADSSKLTIGEPAIAVGSP 193
++ G + +DL+++ + H+ D ++ A++ LTIG + +G P+
Sbjct: 54 SNFTGVEIQSDLSSIIPQTLNHIADQSSVANTRVLTIGVSGL-IGGPI 99

>ref|NP_735272.1| Unknown [Streptococcus agalactiae NEM316]
Length = 383

Score = 26.9 bits (58), Expect = 7.0
Identities = 13/26 (50%), Positives = 17/26 (65%)

Query: 2 AKANIGKLLLTGVVGGAIALGGSAIY 27
A NIGK L T +VG +A+G + IY
Sbjct: 141 ASLNIGKALTTFIVGLVLAIGVNYIY 166

>ref|NP_786635.1| extracellular protein [Lactobacillus plantarum WCFS1]
Length = 322

Score = 26.9 bits (58), Expect = 7.0
Identities = 38/158 (24%), Positives = 65/158 (41%), Gaps = 14/158 (8%)

Query: 140 SGQQKVAKASVVGYDEYTDLAVLKISSEHVKDVTFA---DSSKLTIGEPAIAVGSPLSQ 196
S G + ++ Y+ ++ I++ K+ A A D++ I L
Sbjct: 104 SSGSGINVKILNYNGSNNITT--ITANQYKNAALTAGITDANIYVTSATPIDGSGALAGV 161

Query: 197 FANTATEGILSATSRQVTLTQENGQT---TNINAIQ---TDAAINPGNSGGALINIEGQ 249
+A A G S ++QVT Q+ T T N + TD+ +N +G A + +
Sbjct: 162 YAAYAKSGN-SLNTKQVTAQQDELSTLSGITQANKSKDGYTDSQLNNNAVAG-AKKEMAQK 219

Query: 250 VIGITQSKITTTEGSTSVEGLGFAIPSNDVNNIINKL 287
IT+++ITT + + L I +N IIN L
Sbjct: 220 GSNITKNEITTIIVNQQITNNNLNVITNNQKTEIIINLL 257

>ref|ZP_00046678.1| COG4653: Predicted phage phi-C31 gp36 major capsid-like protein [Lactobacillus gasseri]
Length = 392

ANNEX 4

Score = 26.9 bits (58), Expect = 7.0
Identities = 16/64 (25%), Positives = 28/64 (43%), Gaps = 11/64 (17%)

Query: 140 SGGQKVAKASVVGYDEYTDL-----AVLKISSEHVKDVTAFADSSKLTIGEPAIA 188
+G KA + +D+ DL AV ++ + VK + D + I +P++
Sbjct: 248 AGSTAAKADALTFFDDLIIDLFYSLKAPYRQNAVFLMNDDTVKAIRKMKDKNDQYIWQPSVQ 307

Query: 189 VGSP 192
VG P
Sbjct: 308 VGQP 311

>ref|ZP_00046947.1| COG2931: RTX toxins and related Ca2+-binding proteins
[Lactobacillus
gasseri]
Length = 1991

Score = 26.9 bits (58), Expect = 7.0
Identities = 30/113 (26%), Positives = 43/113 (38%), Gaps = 7/113 (6%)

Query: 114 KKSGGDAYVVTNYHVIAGNSSLDVLLSGGQKVKA
S V G + T V AG L + S ++VKA YD + + V +AT
Sbjct: 1066 KDADGNVYAMTGNPVNAGTYYLHLTKSAIEQVKADNSNYDFTSVNGEFTYTINAVNGIAT 1125

Query: 174 FADSSKLTIGEPAIA--VGSP
+ SS T A+ VS G N G +S Q T + G T
Sbjct: 1126 LSGSSSKTYDGQAVTTAEVNSTNGDIIVNFTFPG---SSAQSTYVLQTGDYT 1174

>ref|ZP_00046780.1| COG3210: Large exoproteins involved in heme utilization or adhesion
[Lactobacillus gasseri]
Length = 3692

Score = 26.9 bits (58), Expect = 7.0
Identities = 28/112 (25%), Positives = 45/112 (40%), Gaps = 7/112 (6%)

Query: 204 GILSATSRQVTLTQENGQTTNINAIQTDAAINPGNSGGALINIEGQVIGITQS
G L +T +NG+ T + Q A+ + G I+ I QS T + D
Sbjct: 2225 GNLTVTDEDGNITSQNGKITWNHESQEFEAVPAIDHDGYYIS----SINQSNSTASVD 2278

Query: 264 GSTSVEGLGFAIPSNDVVNIINKLEADGKISRPA
G T G P++ NI+ L + + A G I +D + T +S+
Sbjct: 2279 GQTGAVGTETVTPNSQNGNIVITLTRNPDV
PVAAGQSINYIDDGQTIESA 2330

>ref|ZP_00062638.1| COG4932: Predicted outer membrane protein [Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293]
Length = 508

Score = 26.6 bits (57), Expect = 9.1
Identities = 33/138 (23%), Positives = 50/138 (36%), Gaps = 29/138 (21%)

Query: 115 KSGGDAYVVTNYHVIAGNSSLD-----VLLSGGQKVKA
KSG D+ + ++ HV N D V L G KV + G L
Sbjct: 171 KSGSDSEINSDVHVYPKNEQTAIKDLSDESKKDLIVTLPDGSKVYNATY
GQKFGYQLQ 230

Query: 160 VLKISSEHVKDVTAFADSSKLTIGEPAIAV--GSPLGSQFANTATEGILSATS
+ + KD D+ L I + A V G G+ + +AT+ T
Sbjct: 231 IAVPWNIADKDTFNVVDTPNLGIDDDATTVKVAGLTKGTDYTVSATDA-----T 279

ANNEX 4

Query: 217 QENGQTTNINAIQTDAAI 234
+NG++ I T AA+
Sbjct: 280 DKNGKSFKITFNPTAAAV 297

>ref|NP_687888.1| exonuclease RexA [Streptococcus agalactiae 2603V/R]
Length = 1207

Score = 26.6 bits (57), Expect = 9.1
Identities = 21/85 (24%), Positives = 33/85 (38%), Gaps = 8/85 (9%)

Query: 292 KISRPALGIRMVSDLSQLSTNDSSQLKLPXXXXXXXXXXXXXXPAASAGLKAGDVITKV 351
KI P L I VD+ + T S KLP A+ G +++ ++
Sbjct: 1010 KIYEPILDIEGVDMETITKTSVDFKLPDFSTSQQ-----DPAALGSAVHELMQRI 1061

Query: 352 GDTAVTSSTDLQSALYSHNINDTVK 376
++ D+Q AL N +VK
Sbjct: 1062 EMSSHVKMEDIQKALTEVNAETSVK 1086

>ref|NP_687383.1| 3-oxoacyl-(acyl-carrier-protein) synthase II [Streptococcus agalactiae 2603V/R]
ref|NP_734805.1| Unknown [Streptococcus agalactiae NEM316]
Length = 410

Score = 26.6 bits (57), Expect = 9.1
Identities = 30/148 (20%), Positives = 58/148 (39%), Gaps = 5/148 (3%)

Query: 117 GGDAYVVTNYHVIAGNSSLVDLLSGGQKVKASVVGYDEYTDLAVLKISSEHVKDVFAD 176
GG +T + IAG SL L + +AS+ + + + S V + A+
Sbjct: 189 GGAEAAITKF-AIAGFQSILTALEPSRASIPFDKDRNGFIMGEGSGMLVLESLEHAE 247

Query: 177 SSKLTIKEPAIAVGSPPLGS-QFANTATEGILSATSRQVTLTQENGQTTNINAIQTDAAIN 235
TI + G+ + + EG+ + + Q+ L + N + +N +
Sbjct: 248 KRGATILAEVVGYGNNTCDAYHMTSPHPEGLGATKAIQLALVEANIKPEEVNYVNAHGTT 307

Query: 236 PGNSGG---ALINIEGQVIGITQSKITT 260
P N G A++ G + ++ +K T
Sbjct: 308 PANEKGESQAIVAAALGTDVPVSSTKSFT 335

Database: Unfinished Lactobacillus gasseri; Completed Lactobacillus plantarum WCFS1; Completed Lactococcus lactis subsp. lactis; Unfinished Leuconostoc mesenteroides subsp. mesenteroides ATCC 8293; Unfinished Oenococcus oeni MCW; Completed Streptococcus agalactiae 2603V/R; Completed Streptococcus agalactiae NEM316

Posted date: Oct 29, 2003 1:28 AM

Number of letters in database: 4,501,851

Number of sequences in database: 15,229

Lambda K H
0.308 0.128 0.338

Gapped
Lambda K H
0.267 0.0410 0.140

Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 404,382
Number of Sequences: 15229
Number of extensions: 14371

ANNEX 4

Number of successful extensions: 43
Number of sequences better than 10.0: 10
Number of HSP's better than 10.0 without gapping: 4
Number of HSP's successfully gapped in prelim test: 6
Number of HSP's that attempted gapping in prelim test: 37
Number of HSP's gapped (non-prelim): 11
length of query: 408
length of database: 4,217,779
effective HSP length: 94
effective length of query: 314
effective length of database: 2,884,107
effective search space: 905609598
effective search space used: 905609598
T: 11
A: 40
X1: 16 (7.1 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 42 (21.7 bits)
S2: 57 (26.6 bits)

ANNEX 5

BLASTP 2.2.6 [Apr-09-2003]

RID: 1065207448-6660-583997.BLASTQ3

Query=

(408 letters)

Database: Completed Streptococcus mutans UA159;

1,531,058 sequences; 495,743,110 total letters

Taxonomy reports

Sequences producing significant alignments:		Score	E	
		(bits)	Value	
ref NP_722446.1	serine protease HtrA [Streptococcus mutans...	296	5e-82	
ref NP_722143.1	putative transcriptional regulator [Strept...	28	0.36	
ref NP_721869.1	putative UDP-N-acetylglucosamine 1-carboxy...	28	0.62	
ref NP_721706.1	putative bacitracin synthetase [Streptococ...	27	0.81	
ref NP_720929.1	putative polyribonucleotide nucleotidyltra...	27	1.1	
ref NP_721524.1	putative ABC transporter, phosphate-bindin...	27	1.4	
ref NP_722399.1	glucan-binding protein A, GbpA [Streptococ...	26	1.8	
ref NP_721986.1	putative D-3-phosphoglycerate dehydrogenas...	26	2.4	
ref NP_722066.1	putative 3-oxoacyl-(acyl-carrier-protein) ...	26	2.4	
ref NP_722435.1	conserved hypothetical protein [Streptococ...	25	3.1	
ref NP_721103.1	phosphoenolpyruvate:sugar phosphotransfера...	24	9.0	
ref NP_720786.1	hypothetical protein [Streptococcus mutans...	24	9.0	

Alignments

>ref|NP_722446.1| serine protease HtrA [Streptococcus mutans UA159]
Length = 402

Score = 296 bits (759), Expect = 5e-82
Identities = 153/286 (53%), Positives = 203/286 (70%), Gaps = 3/286 (1%)

Query: 107 EGSGVIYKKSGGDAYVVTNYHVVIAGNSSLVLLSGGQKVKASVVGYDEYTDLAVLKISSE 166
EGSGVIYKK G AY+VITN HV+ L+++++ G+KV +VG D Y+DLAV+KISS+

ANNEX 5

Sbjct: 107 EGSGVIYKKDGDSAYLVTNNHVVKDAEKLEIMMANGKKVVGKLVGSDTYSDELAVIKISSK 166
Query: 167 HVKDVTAFADSSKLTIGEPAIAVGSPPLGSQFANTATEGILSATSQRQVTLTQENGQTTNIN 226
+V VA FA+S K+ +GEPAIA+GSPPLGS +AN+ TEGI+S+ SR VT ENG+T + N
Sbjct: 167 YVTTVAEFANSNDKIKVGEPAIAIGSPLGSDYANSVTEGIVSSLRTVTSQNENGETISTN 226
Query: 227 AIQTDAAINPGNSGGALINIEGQVIGITQSKITTEDGST--SVEGLGFAIPSNDVVNII 284
AIQTDAAINPGNSGGALINIEGQVIGITQSKI ++ + ++ +VEG+GFAIPSNDVV+II
Sbjct: 227 AIQTDAAINPGNSGGALINIEGQVIGINSSKIASSNNNSNSGVAVEGMGFAIPSNDVVSI 286
Query: 285 NKLEADGKISRPALEGIRMVQLSQLSTNDSSQLKLPXXXXXXXXXXXXXXLPAASAGLKA 344
N+LE +G++ RPALGI M +LS+ ST+ LK+P +P A LK
Sbjct: 287 NQLEENGEVVRPALGISMANLSEASTSGRDTLKIPSDVTSGIVVLSTQSGMP-ADGKLKK 345
Query: 345 GDVITKVGDATVTSSTDQLQSALYSHNINDTVKVTYYRDGKSNTADV 390
DVIT++ V S +DLQS LY H D +K+T+YR+ T ++
Sbjct: 346 YDVITEIDGKKVASISDLQSILYKHKGDKIKLTFYREKDKQTVEI 391

>ref|NP_722143.1| putative transcriptional regulator [Streptococcus mutans UA159]

Length = 261

Score = 28.5 bits (62), Expect = 0.36
Identities = 34/128 (26%), Positives = 51/128 (39%), Gaps = 26/128 (20%)

Query: 193 LGSQFANTATEGILSATSQRQVTLTQENGQTTNINAIQTDAAIN--PGNSGGALINIEGQ 249
+G Q N TE L T R+ T T + + ++ AAI GN G ++
Sbjct: 154 VGIQLLNLQTEN-LEETIRKQTAINMAINTLSYSEMKAVALNELDGNEGRLTASVIAD 212
Query: 250 VIGITQSKITTEDGSTSVEGLGFAIPSNDVVNIIINKLEADGKISRPALEGIRMVQLSQLS 309
IGIT+S I VN + KLE+ G I +LG++ L ++
Sbjct: 213 RIGITRSVI-----VNALRKLESAGIIIESRSLGMKGTYLKVIN 250
Query: 310 TNDSSQLK 317
+LK
Sbjct: 251 EGIFDKLK 258

>ref|NP_721869.1| putative UDP-N-acetylglucosamine 1-carboxyvinyltransferase [Streptococcus mutans UA159]

Length = 423

Score = 27.7 bits (60), Expect = 0.62
Identities = 34/146 (23%), Positives = 64/146 (43%), Gaps = 10/146 (6%)

Query: 150 VGYDEYTDLVLKISSEHVKDVATFADSSKLTIGEPAIAVGSPPLGSQ--FANTATEGILS 207
V +DE + ++ + + + DVA + S++ +I V P+ ++ A + G +
Sbjct: 66 VDFDEERNQILVDAVDGD-ILDVAPYEYVSQM---RASIVVLPILARNGHAKVSMPPGCT 121
Query: 208 ATSRQVTLTQENGQTTNINAIQT--DAAINPGNSGGALINIEGQVIGITQSKI--TTTED 263
SR + L + + QT D GA I ++ +G TQ+ + T D
Sbjct: 122 IGSRPIDLHLKGLEAMGAKIQQTGGDITATADRLKGANIYMDFPSVGATQNLMMMAATLAD 181
Query: 264 GSTSVEGLGFAIPSNDVVNIINKLEA 289
G+T +E D+ N++NK+ A
Sbjct: 182 GTTIIENAAREPEIVDLANLLNKMGA 207

ANNEX 5

>ref|NP_721706.1| putative bacitracin synthetase [Streptococcus mutans UA159]
Length = 1455

Score = 27.3 bits (59), Expect = 0.81
Identities = 20/63 (31%), Positives = 29/63 (46%), Gaps = 3/63 (4%)

Query: 119 DAYVVTNYHVIAGNSSLDVLLSGGQKVKASVVGYDEYT-DIALVLKISSEHVKDVFATFADS 177
D V Y V G +D ++S K+K + Y E T +L S EH D+ T +
Sbjct: 887 DQVKVNGYRVELGE--IDSIISKMSKIKRAKTIYQEETGNLIAFCESKEHCSDIETRKEL 944

Query: 178 SKL 180
SK+
Sbjct: 945 SKI 947

>ref|NP_720929.1| putative polyribonucleotide nucleotidyltransferase (general stress
protein 13) [Streptococcus mutans UA159]

Length = 130

Score = 26.9 bits (58), Expect = 1.1
Identities = 13/25 (52%), Positives = 16/25 (64%)

Query: 139 LSQQQKVKAQVVGYDEYTDALVLI 163
LS GQ+V VV YDEY+ A L +
Sbjct: 50 LSVGQEVLVQVVDYDEYSQKASL 74

>ref|NP_721524.1| putative ABC transporter, phosphate-binding protein
[Streptococcus

mutans UA159]
Length = 287

Score = 26.6 bits (57), Expect = 1.4
Identities = 17/44 (38%), Positives = 21/44 (47%), Gaps = 1/44 (2%)

Query: 335 LPAASAGLKAGDVITKVGDTAVTSSTDLQSALYSH-NINDTVKV 377
L A S + G IT VG TA+ + S + H NI TV V
Sbjct: 19 LAACSNWIDKGQSITSVGSTALQPLVEASSDEFGHANIGKTVNV 62

>ref|NP_722399.1| glucan-binding protein A, GbpA [Streptococcus mutans UA159]
Length = 565

Score = 26.2 bits (56), Expect = 1.8
Identities = 30/153 (19%), Positives = 54/153 (35%), Gaps = 12/153 (7%)

Query: 172 ATFAADSSKLTIGEPAIAVGSPPLGSQFANTATEGILSATSRQVTLTQENGQTTNINAIQTD 231
AT +SS+ + E A S N + + S+ + + + G+ + A+
Sbjct: 55 ATVQESSEQPVTEAPAA----DSSVENNSANAVKSSETAEAAEVSDGGRASQTEAVTNQ 109

ANNEX 5

Query: 232 AAINPGNSGGALINIEGQVIGITQSKitTEDGSTSVEGLGFAIPSNDVVNIINKLEADG 291
+ + G+ + + T+ A +ND + E DG
Sbjct: 110 TNSEHHHPAEKATAVSGEAQSVQNAPSENAAQQETAKTEPATAAENNDAAPTNSFFEKDG 169

Query: 292 K-----ISRPALGIRMVSQLSTN-DSSQLK 317
K + A G + +D QL N D SQ+K
Sbjct: 170 KWYYKKADGQLATGWQTIDGKQLYFNQDGSQLVK 202

>ref|NP_721986.1| putative D-3-phosphoglycerate dehydrogenase [Streptococcus mutans
UA159]
Length = 393

Score = 25.8 bits (55), Expect = 2.4
Identities = 17/78 (21%), Positives = 32/78 (41%), Gaps = 3/78 (3%)

Query: 138 LLSGGQKVKA SVVGYDEYTDLAVLKISSEHVKDVATFADSSKLTIGEPAIAVGSPGSQF 197
+ + +++ +V+GYD Y + S HVK V D + I + PL +
Sbjct: 151 IANDARRLGMNVLGDPYVSIETAWNISSHVKRVNDLKD---IFENSODYITIHVPLNDT 207

Query: 198 ANTATEGILSATSRQVTL 215
NT + ++ T+
Sbjct: 208 KNTFNADSFALMNKGTTI 225

>ref|NP_722066.1| putative 3-oxoacyl-(acyl-carrier-protein) synthase [Streptococcus mutans
UA159]
Length = 410

Score = 25.8 bits (55), Expect = 2.4
Identities = 27/136 (19%), Positives = 48/136 (35%), Gaps = 13/136 (9%)

Query: 136 DVLLSGGQKVKA SVVGYDEYTDLAVLKISSEHVKDVATF-ADSSKLTIGEPAIAVGSPLG 194
DV+L+GG + + +G + L L + + + F D + +GE A
Sbjct: 184 DVILAGGSEASITKIGIGGFNALTALSTTEDPARSAIPFDKDRNGFVMGEGA----- 235

Query: 195 SQFANTATEGILSATSRQVTLTQE-NGQTTNINAIQTDAAINPGNSGGALINIEGQVIGI 253
A E + A R + E G +N +A G+ I + GI
Sbjct: 236 ---AVLILESLEHAQKRGARILAEVVGYSNCDAYHMTTPDGSGAAKAIKLAINEAGI 292

Query: 254 TQSKITTEDGSTSVE 269
+ ++ TS +
Sbjct: 293 SPEEVNYVNAHGTSTQ 308

>ref|NP_722435.1| conserved hypothetical protein [Streptococcus mutans UA159]
Length = 325

Score = 25.4 bits (54), Expect = 3.1
Identities = 18/61 (29%), Positives = 27/61 (44%), Gaps = 12/61 (19%)

Query: 198 ANTATEG----ILSATSRQVTLTQENGOTTNINAIQTDAAINPGNSGGALINIEGQVIGI 253
A T+G L+AT + T+T G TT++ + G+ G I I GQ +

ANNEX 5

Sbjct: 258 ATNTTDGESGTTLTATDKTYTVTLAEGSTTSM-----LTVGSPSGVEITINGQKVDT 309

Query: 254 T 254

T

Sbjct: 310 T 310

>ref|NP_721103.1| phosphoenolpyruvate:sugar phosphotransferase system enzyme I, PTS

system EI component [Streptococcus mutans UA159]

Length = 577

Score = 23.9 bits (50), Expect = 9.0

Identities = 19/104 (18%), Positives = 39/104 (37%), Gaps = 3/104 (2%)

Query: 210 SRQVTLTQENGQTTNINAIQTDAAINPGNSGGALINIEGVIGITQSKitTEDGST-SV 268

+ +T +NG +N I INP A G+ +++ +D T +

Sbjct: 207 TNDITERVKNGDIVAVNGITGQVIINPTEDQIAEFKAAGETYAKQKAEWALLKDAETVTA 266

Query: 269 EGLGFAIPSNDVVNIINKLEADGKISRPAVGIRMVQLSQLSTND 312

+G F + +N + +E A+G+ + + + D

Sbjct: 267 DGKHFEALAN--IGTPKDVEGVNNNGAEAVGLYRTEFLYMQDSQD 308

>ref|NP_720786.1| hypothetical protein [Streptococcus mutans UA159]

Length = 411

Score = 23.9 bits (50), Expect = 9.0

Identities = 12/40 (30%), Positives = 19/40 (47%)

Query: 151 GYDEYTDLAVLKISSEHVKDVTAFADSSKLIGEPAAVG 190

GY +T + +S + K T AD LT+G+ + G

Sbjct: 301 GYAYFTSKDIKTVSEKSYKSDWTQADVDALTGVGDSSTGKG 340

Database: Completed Streptococcus mutans UA159

Posted date: Oct 1, 2003 10:43 PM

Number of letters in database: 579,702

Number of sequences in database: 1960

Lambda	K	H
0.308	0.128	0.338

Gapped

Lambda	K	H
0.267	0.0410	0.140

Matrix: BLOSUM62

Gap Penalties: Existence: 11, Extension: 1

Number of Hits to DB: 35,974

Number of Sequences: 1960

Number of extensions: 1254

ANNEX 5

Number of successful extensions: 5
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 0
Number of HSP's successfully gapped in prelim test: 1
Number of HSP's that attempted gapping in prelim test: 5
Number of HSP's gapped (non-prelim): 1
length of query: 408
length of database: 577,947
effective HSP length: 81
effective length of query: 327
effective length of database: 419,835
effective search space: 137286045
effective search space used: 137286045
T: 11
A: 40
X1: 16 (7.1 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 42 (21.7 bits)
S2: 50 (23.9 bits)

ANNEX 6

BLASTP 2.2.6 [Apr-09-2003]

RID: 1065207742-9645-69613.BLASTQ3

Query=

(408 letters)

Database: Completed Streptococcus pneumoniae R6;

1,531,058 sequences; 495,743,110 total letters

Taxonomy reports

Score (bits)	E Value
Sequences producing significant alignments:	
325	2e-90
ref NP_359636.1 Serine protease [Streptococcus pneumoniae R6]	
40	2e-04
ref NP_359374.1 Conserved hypothetical protein [Streptococ...	
29	0.22
ref NP_357681.1 Hypothetical protein [Streptococcus pneumo...	
28	0.48
ref NP_357856.1 Alcohol dehydrogenase, propanol-preferring...	
27	1.4
ref NP_357669.1 Conserved hypothetical protein [Streptococ...	
25	3.1
ref NP_358044.1 EcoA type I restriction-modification enzym...	
25	5.3
ref NP_359110.1 Penicillin-binding protein 2B [Streptococc...	
24	6.9
ref NP_358390.1 6-phosphofructokinase I [Streptococcus pne...	
24	6.9
ref NP_359031.1 Isochorismatase [Streptococcus pneumoniae R6]	
24	9.0
ref NP_359032.1 Transcriptional pleiotropic repressor [Str...	

Alignments

>ref|NP_359636.1| Serine protease [Streptococcus pneumoniae R6]
Length = 397

Score = 325 bits (832), Expect = 2e-90
Identities = 167/284 (58%), Positives = 212/284 (74%), Gaps = 3/284 (1%)

Query: 107 EGSGVIYKKSGGDAYVVTNYHVIAGNSSLVDVLLSGGQKVKASVVGYDEYTDLAVLKISSE 166
EGSGVIYKK+ +AY+VTN+HVI G S +D+ LS G KV +VG D ++D+AV+KISSE
Sbjct: 106 EGSGVIYKKNDKEAYIVTNHHVINGASKVDIRLSGDTKVPGEIVGADTFSDIAVVKISSE 165

ANNEX 6

Query: 167 HVKDVTAFADSSKLTIGEPAIAVGSPPLGSQFANTATEGILSATSRSRQVTLTQENGQTTNNIN 226
V VA F DSSKLT+GE AIA+GSPLGS++ANT T+GI+S+ +R V+L E+GQ +
Sbjct: 166 KVTTVAEFGDSSKLTVGETAIAIGSPLGSEYANTVTQGIVSSLNRNVSLKSEDGQAISTK 225

Query: 227 AIQTDAAINPGNSGGALINIEGQVIGITQSKITTEDGSTSVEGLGFAIPSNDVVNIINK 286
AIQTD AINPGNSGGALINIEGQVIGIT SKI T +G TSVEGLGFAIP+ND +NII +
Sbjct: 226 AIQTDAAINPGNSGGPLINIEGQVIGITSSKIAT--NGGTSVEGLGFAIPPANDAINIIEQ 283

Query: 287 LEADGKISRPAALGIRMVQLSQLSTNDSSQLKLPPXXXXXXXXXXXXXXLPAASAGLKAGD 346
LE +GK+ +RPALGI+MV+LS +ST+D +L +P +P A+ L+ D
Sbjct: 284 LEKNGKVTRPALGIQMVNLSNVSTDIRRNLNIPSNVTSGVIVRSVQSNMP-ANGHLEKYD 342

Query: 347 VITKVGDTAVTSSTDQLQSALYSHNINDTVKVTYYRDGKSNTADV 390
VITK V D + SSTDQLQSALY+H+I DT+K+TYYR+GK T +
Sbjct: 343 VITKVDDEKIASSTDQLQSALYNHSIGDTIKITYYRNGKEETTSI 386

>ref|NP_359374.1| Conserved hypothetical protein [Streptococcus pneumoniae R6]
Length = 345

Score = 39.7 bits (91), Expect = 2e-04
Identities = 19/48 (39%), Positives = 27/48 (56%)

Query: 342 LKAGDVITKVGDTAVTSSTDQLQSALYSHNINDTVKVTYYRDGKSNTAD 389
L D +T V D SS DL + S + D+VKVTY DG++ +A+
Sbjct: 146 LNISDTVTAVNDQTFDSSKDLIDYVSSQKLGDSVKVTYEEDGQTKSAE 193

>ref|NP_357681.1| Hypothetical protein [Streptococcus pneumoniae R6]
Length = 320

Score = 29.3 bits (64), Expect = 0.22
Identities = 21/72 (29%), Positives = 33/72 (45%), Gaps = 5/72 (6%)

Query: 189 VGSPLGSQFANTATEGILSATSRSRQVTLTQENGQTTNNINAIQTDAAINPGNSGGALINIEG 248
+ S S+F T + + ++ GQTT INA +A + N+ ++ IEG
Sbjct: 142 IASSYSSRFEEVILRLPKGRRTLGINISANRGQTTIINASLENATL--NTNSYILRIEG 198

Query: 249 QVIGITQSKITT 260
I SK+TT
Sbjct: 199 S--RIKNASKLTT 208

>ref|NP_357856.1| Alcohol dehydrogenase, propanol-preferring. [Streptococcus pneumoniae R6]
Length = 339

Score = 28.1 bits (61), Expect = 0.48
Identities = 18/73 (24%), Positives = 40/73 (54%), Gaps = 4/73 (5%)

Query: 109 SGVIYKKSGGDAY--VVTNYHVIAGNSSLDVLLSGGQKVKA SVVGYDEYTDLAVLKISSE 166
+G+I +K+ G A+ VVT +A N ++D + +GG+ V + E +L+++K +
Sbjct: 221 AGLIKEKTDGGAHSAVVTAVSKVAFNQAVDSIRAGGRVVAVGLP--SEMELLSIVKTVID 278

Query: 167 HVKDVTAFADSSK 179

ANNEX 6

++ + + + K
Sbjct: 279 GIQVIGSLVGTRK 291

>ref|NP_357669.1| Conserved hypothetical protein [Streptococcus pneumoniae R6]
Length = 1161

Score = 26.6 bits (57), Expect = 1.4
Identities = 11/25 (44%), Positives = 16/25 (64%)

Query: 335 LPAASAGLKAGDVITKVGDTAVTSS 359
LP ++G K GD+ K GDT +T +
Sbjct: 53 LPEETSGTKEGDLSEKPGDTVLTQA 77

>ref|NP_358044.1| EcoA type I restriction-modification enzyme R subunit
[Streptococcus pneumoniae R6]
Length = 777

Score = 25.4 bits (54), Expect = 3.1
Identities = 18/66 (27%), Positives = 30/66 (45%), Gaps = 1/66 (1%)

Query: 119 DAYVVTNYHVIAGNSSLDVLLSGGQKVKA\$VVGYDEYTDLAVLKISSEHVKDVA\$FADSS 178
+ Y+VT+ V NS++ VL G+ + S+ Y L ++ + V AD
Sbjct: 588 EKYIVTDKQVTLN\$TVQVL\$ENGKLITESLTDYTRKNILGSYATLNDFI-TVWHTADKK 646

Query: 179 KLTIGE 184
KL + E
Sbjct: 647 KLILDE 652

>ref|NP_359110.1| Penicillin-binding protein 2B [Streptococcus pneumoniae R6]
Length = 685

Score = 24.6 bits (52), Expect = 5.3
Identities = 14/47 (29%), Positives = 19/47 (40%)

Query: 190 GSPLGSQFANTATEGILSATSRQVTLTQENGQT\$TNINAIQTDA\$AINP 236
G G F+N A I T + + Q TN NA+ + NP
Sbjct: 602 GLTTGRAFSNGALVSISGKTGAESYVADGQQATNTNAVAYAPS\$DNP 648

>ref|NP_358390.1| 6-phosphofructokinase I [Streptococcus pneumoniae R6]
Length = 335

Score = 24.3 bits (51), Expect = 6.9
Identities = 14/49 (28%), Positives = 24/49 (48%)

Query: 246 IEGQVIGITQS\$KITT\$EDGSTVEGLGFAIPSNDVVNIINKLEADGKIS 294
I G +GI K+ T+ EG F++ + + + N EAD ++S
Sbjct: 280 IGGVAVGIRNEKMVENPILGTAAEGALFSLTAEGKIVVNNPHEADIELS 328

ANNEX 6

>ref|NP_359031.1| Isochorismatase [Streptococcus pneumoniae R6]
Length = 191

Score = 24.3 bits (51), Expect = 6.9
Identities = 13/31 (41%), Positives = 18/31 (58%), Gaps = 2/31 (6%)

Query: 163 ISSEHVKDVATFADSSKLTIGEPAIAVGSP 193
IS ++ +D ADS KLT G PA A+ +
Sbjct: 6 ISIDYTEDFV--ADSGKLTAGAPAQAISDAI 34

>ref|NP_359032.1| Transcriptional pleiotropic repressor [Streptococcus pneumoniae R6]
Length = 262

Score = 23.9 bits (50), Expect = 9.0
Identities = 24/105 (22%), Positives = 40/105 (38%), Gaps = 22/105 (20%)

Query: 213 VTLTQENGQTTNINAIQTDAAINPGNSGGALINIEGQVIGITQSKitTEDGSTSVEGLG 272
VT+ + + A+ GN G ++ IGIT+S I
Sbjct: 176 VTMAVNTLSYSELRAVSAILGELNGNEGQLTASVIADRIGITRSVI----- 221

Query: 273 FAIPSNDVVNIINKLEADGKISRPAALGIRMVQLSQLSTNDSSQLK 317
VN + KLE+ G I +LG++ L L ++ ++K
Sbjct: 222 -----VNALRKLESAGIIESRSLGMKGTYLKVLISDIFEVK 258

Database: Completed Streptococcus pneumoniae R6
Posted date: Oct 1, 2003 10:43 PM
Number of letters in database: 589,192
Number of sequences in database: 2043

Lambda K H
0.308 0.128 0.338

Gapped
Lambda K H
0.267 0.0410 0.140

Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 33,145
Number of Sequences: 2043
Number of extensions: 1151
Number of successful extensions: 4
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 0
Number of HSP's successfully gapped in prelim test: 1
Number of HSP's that attempted gapping in prelim test: 4
Number of HSP's gapped (non-prelim): 1
length of query: 408
length of database: 588,593

ANNEX 6

```
effective HSP length: 81
effective length of query: 327
effective length of database: 423,515
effective search space: 138489405
effective search space used: 138489405
T: 11
A: 40
X1: 16 ( 7.1 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 42 (21.7 bits)
S2: 50 (23.9 bits)
```

ANNEX 7

BLASTP 2.2.6 [Apr-09-2003]

RID: 1065207829-10461-370696.BLASTQ3

Query=

(408 letters)

Database: Completed Streptococcus pyogenes SSI-1;

1,531,058 sequences; 495,743,110 total letters

Taxonomy reports

Sequences producing significant alignments:		Score	E	
		(bits)	Value	
ref NP_803122.1	putative serine protease [Streptococcus py...	311	2e-86	
ref NP_801300.1	putative alcohol dehydrogenase I [Streptoc...	32	0.030	
ref NP_802185.1	putative phage-related tail protein [Strept...	31	0.067	
ref NP_802392.1	putative tail protein, phage associated [S...	29	0.20	
ref NP_801584.1	putative transcriptional pleiotropic repre...	29	0.26	
ref NP_803013.1	putative transcriptional regulator [Streptoc...	25	2.8	
ref NP_802639.1	conserved hypothetical protein [Streptococ...	25	4.8	
ref NP_801305.1	50S ribosomal protein L4 [Streptococcus py...	25	4.8	

Alignments

>ref|NP_803122.1| putative serine protease [Streptococcus pyogenes SSI-1]
Length = 407

Score = 311 bits (797), Expect = 2e-86
Identities = 163/288 (56%), Positives = 207/288 (71%), Gaps = 5/288 (1%)

Query: 107 EGSGVIYKKSGGDAYVVTNYHVIAGNSSLVDLLSGGQKVAKASVVGYDEYTDIAVLKISSE 166
EGSGVIY+K G AYVVTN HVI G +++L++ G KV +VG D Y+DLAV+KISS+
Sbjct: 108 EGSGVIYRKDGNSAYVVTNNHVIDGAKRIEILMADGSKVVGELVGADTYSDIAVVKISSD 167

Query: 167 HVKDVAFTADSSKLTIGEPAIAVGSPPLGSQFANTATEGILSATSRQVTLTQENGQTTNIN 226
+K VA FADS+KL +GE AIA+GSPLG+Q+AN+ T+GI+S+ SR VTL ENG+T + N

ANNEX 7

Sbjct: 168 KIKTVAEFADSTKLNVGEVAIAIGSPLGTQYANSVTQGIVSSLSRTVLKNENGETVSTN 227
Query: 227 AIQTDAAINPGNNSGGALINIEGQVIGITQSKitTTEDGST---SVEGLGFAIPSNDVVN 282
AIQTDAAINPGNNSGGALINIEGQVIGITQSKitTTEDGST---SVEGLGFAIPSNDVVN 282
Sbjct: 228 AIQTDAAINPGNNSGGPLINIEGQVIGINSSKISSTPTGSNGNSGAVEGIGFAIPSTDVIK 287
Query: 283 IINKLEADGKISRPAVGIRMVDLSQLSTNDSQLKLPXXXXXXXXXXXXXPAASAGL 342
II +LE +G++ RPALGI MV+L+ LSTN SQ+ +P +P AS L
Sbjct: 288 IIKQLETNGEVIRPALGISMVNLDLSTNALSQINIPSVTGGIVVAEVKEGMP-ASGKL 346
Query: 343 KAGDVITKVGDATVTSSTDLQSLAYSHNINDTVKVTYRDGKSNTADV 390
DVIT++ V S +DLQS+LY H+INDT+KVT+YR AD+
Sbjct: 347 AQYDVITEIDGKTVNSISDLQSSLYGHIDINDTIKVTFYRGTTKKADI 394

>ref|NP_801300.1| putative alcohol dehydrogenase I [Streptococcus pyogenes SSI-1]
Length = 282

Score = 32.0 bits (71), Expect = 0.030
Identities = 21/71 (29%), Positives = 37/71 (52%), Gaps = 3/71 (4%)

Query: 110 GVIYKKSGG-DAYVVTNYHVIAGNSSLDVLLSGGQKVKA\$VVGYDEYTDLAVLKISSEHV 168
G I +K+GG VVT +A N ++D + +GG V + EY +L+++K + +
Sbjct: 166 GYIQEKTGGAHGVVTAVKVAFNQAIDSVRAGGTVVAVGLP--SEYMELSIKVTLGI 223

Query: 169 KDVATFADSSK 179
K V + + K
Sbjct: 224 KVVGSLVGTRK 234

>ref|NP_802185.1| putative phage-related tail protein [Streptococcus pyogenes SSI-1]
Length = 1307

Score = 30.8 bits (68), Expect = 0.067
Identities = 48/213 (22%), Positives = 86/213 (40%), Gaps = 30/213 (14%)

Query: 153 DEYTDLAVLKISSEHVKDVTADSSKLTIGEPAIAVGSPLSQFANTATEGILSATSRQ 212
DE ++ K+S + ++ +A +S + I A A G T ILS +
Sbjct: 192 DETATVSYAKLS-QGIRQMAKELPASAVEIAHVAEAAGQ-----LGVKTGDILSFSRTM 244

Query: 213 VTLTQENGOTTNINAIQTDAAINPGNSGGALINIEGQVIGITQSKitTTEDGSTSVEGLG 272
+ L G++TN++A + +I + NI G + S+ + ++ G
Sbjct: 245 IDL---GESTNLSAEEAATSI-----AKIANITG---LASSEYSRFGSAVVAL-GNN 289

Query: 273 FAIPSNDVNNIINKLEADGKISRPAVGIRMVDLSQLSTNDSQLKLPXXXXXXXXXXXX 332
FA D+V + N++ A GK++ G+ + + L+T SS + +
Sbjct: 290 FATTEKDIVAMTNRIAASGKLA----GLTNQEMLALATAMSS-VGIEAEAGGTAMTQSL 344

Query: 333 XXLPAASAGLKAGDVITKVGDATVTSSTDLQSA 365
A ++G GD + K A SS D A
Sbjct: 345 AIERAVASG---GDNLNKFAQIANMSSADFARA 374

ANNEX 7

>ref|NP_802392.1| putative tail protein, phage assocaited [Streptococcus pyogenes

SSI-1]

Length = 1372

Score = 29.3 bits (64), Expect = 0.20

Identities = 34/146 (23%), Positives = 57/146 (39%), Gaps = 19/146 (13%)

Query: 220 QOTTNINAIQTDAAINPGNSGGALINIEGQVIGITQSKitTEDGSTSVEGLGFAIPSND 279
GQ+TN+A + ++I + NI G SK + S G F+ D
Sbjct: 254 GQSTNLSAEEAASSI-----AKIANITGLT-----SKEYSRFGSSVVALGNNFSTTERD 302

Query: 280 VVNIINKLEADGKISRPAKGIRMVDSLSQLSTDSSQLKLPXXXXXXXXXXXXXLPAA 339
V+ + N++ A GK++ G+ ++ L+T SS +
Sbjct: 303 VIAMTNRIAASGKLA---GLTNQEMLALATAMSS---VGIEAEAGGTAMTQTLSAIET 354

Query: 340 AGLKAGDVITKVGDITAVTSSTDLSQA 365

A + G+ +TK A SS D A

Sbjct: 355 AVINGGEDLTKFQAQIANMSSKDFAKA 380

>ref|NP_801584.1| putative transcriptional pleiotropic repressor [Streptococcus pyogenes

SSI-1]

Length = 260

Score = 28.9 bits (63), Expect = 0.26

Identities = 34/128 (26%), Positives = 52/128 (40%), Gaps = 26/128 (20%)

Query: 193 LGSQFANTATEGILSATSRQVTLTQENGQTTNNINAIQTDAAIN---PGNSGGALINIEGQ 249
+G Q N TE L T R+ T T + + ++ AAI GN G ++
Sbjct: 154 VGIQLLNLTEN-LEDTIRKQTAVNMAINTLSYSEMKAVALGELDGNEGRLTASVIAD 212

Query: 250 VIGITQSKitTEDGSTSVEGLGFAIPSNDVVNIINKLEADGKISRPAKGIRMVDSLSQLS 309
I GIT+S I VN + KLE+ G I +LG++ L ++
Sbjct: 213 RIGITRSVI-----VNALRKLESAGIIESRSLGMKGTYLKVIN 250

Query: 310 TNDSSQLK 317

++LK

Sbjct: 251 EGIFAKLK 258

>ref|NP_803013.1| putative transcriptional regulator [Streptococcus pyogenes

SSI-1]

Length = 326

Score = 25.4 bits (54), Expect = 2.8

Identities = 21/79 (26%), Positives = 36/79 (45%), Gaps = 7/79 (8%)

Query: 112 IYKKSGGDAYVVTNYHVIAGNSSLDVLLSGGQK-VKASVVGYDEYTDLAVLKISSEHVKD 170
IY +GG +++ YHV L + G + A V+ D++ +L+ S++ D
Sbjct: 138 IYPLAGGPISHINAKYHVNTLVYRLARIFHGNSAFMNAMVIQEDKHLAKGILQ--SKYFND 195

Query: 171 VATFADSSKLTI---GEP 185

+ T D L + GEP

Sbjct: 196 ILTSWDQQLDLALVGIGGEP 214

ANNEX 7

>ref|NP_802639.1| conserved hypothetical protein [Streptococcus pyogenes SSI-1]
Length = 574

Score = 24.6 bits (52), Expect = 4.8
Identities = 10/32 (31%), Positives = 20/32 (62%)

Query: 345 GDVITKVGDTAVTSSTDLQSALYSHNINDTVK 376
G +I K D+ +TS + + AL++ +ND ++
Sbjct: 23 GVIIKRNDLSLITSLEERKQALFALPVNDEIE 54

>ref|NP_801305.1| 50S ribosomal protein L4 [Streptococcus pyogenes SSI-1]
Length = 207

Score = 24.6 bits (52), Expect = 4.8
Identities = 12/33 (36%), Positives = 18/33 (54%), Gaps = 1/33 (3%)

Query: 262 EDGSTSVEGLGFAIPSN-DVVNIINKLEADGKI 293
ED +VEGL FA P + +++ L D K+
Sbjct: 120 EDKFVAVEGLSFAAPKTAEFAKVLALSIDTKV 152

Database: Completed Streptococcus pyogenes SSI-1

Posted date: Oct 1, 2003 10:43 PM
Number of letters in database: 534,258
Number of sequences in database: 1861

Lambda K H
0.308 0.128 0.338

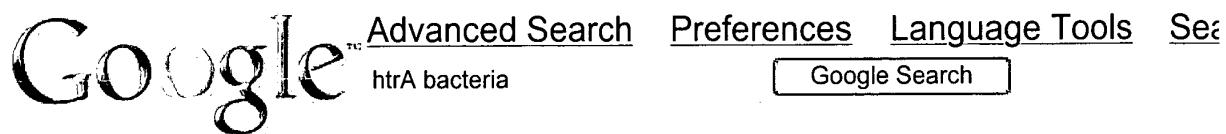
Gapped
Lambda K H
0.267 0.0410 0.140

Matrix: BLOSUM62

Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 32,970
Number of Sequences: 1861
Number of extensions: 1177
Number of successful extensions: 3
Number of sequences better than 10.0: 1
Number of HSP's better than 10.0 without gapping: 0
Number of HSP's successfully gapped in prelim test: 1
Number of HSP's that attempted gapping in prelim test: 3
Number of HSP's gapped (non-prelim): 1
length of query: 408
length of database: 532,687
effective HSP length: 80
effective length of query: 328
effective length of database: 384,447
effective search space: 126098616
effective search space used: 126098616

ANNEX 7

T: 11
A: 40
X1: 16 (7.1 bits)
X2: 38 (14.6 bits)
X3: 64 (24.7 bits)
S1: 42 (21.7 bits)
S2: 50 (23.9 bits)



Web · Images · Groups · Directory · News ·

Searched the web for **htrA bacteria**. Results 1 - 10 of about 1,220. Search took (

Site-directed mutagenesis of the **HtrA** (DegP) serine protease, ...

... An identical effect was observed when His105 was mutated. The mutated

HtrA

were unable to suppress thermosensitivity of the **htrA bacteria**. ...

[www.ncbi.nlm.nih.gov/entrez/ query.fcgi?](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&list_uids=96001242&dopt=Abstract)

[cmd=Retrieve&db=PubMed&list_uids=96001242&dopt=Abstract](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?holding=npg&cmd=Retrieve&db=PubMed&list_uids=9383148&a...) - Similar pages

The **HtrA** family of serine proteases.

... Homologues of **HtrA** have been described in a wide range of bacteria and in eukaryotes.

Its chief role is to degrade misfolded proteins in the periplasm. ...

[www.ncbi.nlm.nih.gov/entrez/ query.fcgi?](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?holding=npg&cmd=Retrieve&db=PubMed&list_uids=9383148&a...)

[holding=npg&cmd=Retrieve&db=PubMed&list_uids=9383148&a... -](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?holding=npg&cmd=Retrieve&db=PubMed&list_uids=9383148&a...)

Similar pages

[More results from www.ncbi.nlm.nih.gov]

Transcriptional Activation of the **htrA** (High-Temperature ...)

... in the ratio of bacterial GFP to DNA was detected for intracellular **bacteria** compared

to extracellular **bacteria**, indicating upregulation of **htrA** upon invasion ...

www.ncbi.nlm.nih.gov/entrez/abstract.cgi?artid=101561 - Similar pages

[PDF] Role of the **htrA** Gene in *Klebsiella pneumoniae* Virulence

File Format: PDF/Adobe Acrobat - [View as HTML](#)

... Alignment of the deduced amino acid sequence of *K. pneumoniae* **HtrA** revealed

extensive amino acid identity with **HtrA** sequences from other **bacteria**. ...

www.ncbi.nlm.nih.gov/entrez/abstract.cgi?artid=101561 - Similar pages

[PDF] Molecular biology of Gram-positive **bacteria**

File Format: PDF/Adobe Acrobat - [View as HTML](#)

... produce small bacteriocidal peptides that can kill closely related **bacteria**, the so ... that processing is carried out by the housekeeping surface protease **HtrA**.

...

www.ncbi.nlm.nih.gov/entrez/abstract.cgi?artid=101561 - Similar pages

Origin and evolution of eukaryotic apoptosis: the bacterial ...
... like protease superfamily, apoptotic ATPases and NACHT family NTPases,
and mitochondrial
HtrA-like proteases, have diverse homologs in **bacteria**, but not in ...
genomebiology.com/pubmed/11965492 - Similar pages

Annual report of Biochimie Microbienne for year 2001
... In *S. aureus*, two **HtrA**-like proteases are present and nothing is known ...
ArlS/arlR regulon
and contribute to the understanding of the virulence of this **bacteria**. ...
www.pasteur.fr/recherche/RAR/RAR2001/Bmic-en.html - 22k - Cached - Similar pages

IAI -- Abstracts: Li et al. 64 (6): 2088
... temperature requirement (**HtrA**) family of stress response proteins are induced by
different environmental stress conditions in a variety of **bacteria** and have ...
iai.asm.org/cgi/content/abstract/64/6/2088 - Similar pages

IAI -- Miller and Shinnick 68 (1): 387
... type **bacteria** (13, 22, 31). The precise role of **HtrA** in intracellular survival of these **bacteria** is not known. *M. tuberculosis* open ...
iai.asm.org/cgi/content/full/68/1/387 - Similar pages
[More results from iai.asm.org]

J. Bact -- Abstracts: Boucher et al. 178 (2): 511
... to **HtrA** (DegP), a serine protease involved in proteolysis of abnormal proteins and required for resistance to oxidative and heat stress in enteric **bacteria**. ...
jb.asm.org/cgi/content/abstract/178/2/511 - Similar pages

Google ►
Result Page: [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) [8](#) [9](#) [10](#) [Next](#)

Google Home - Advertise with Us - Business Solutions - Services & Tools -
Jobs, Press, & Help

©2003 Google